M201 1 and 2

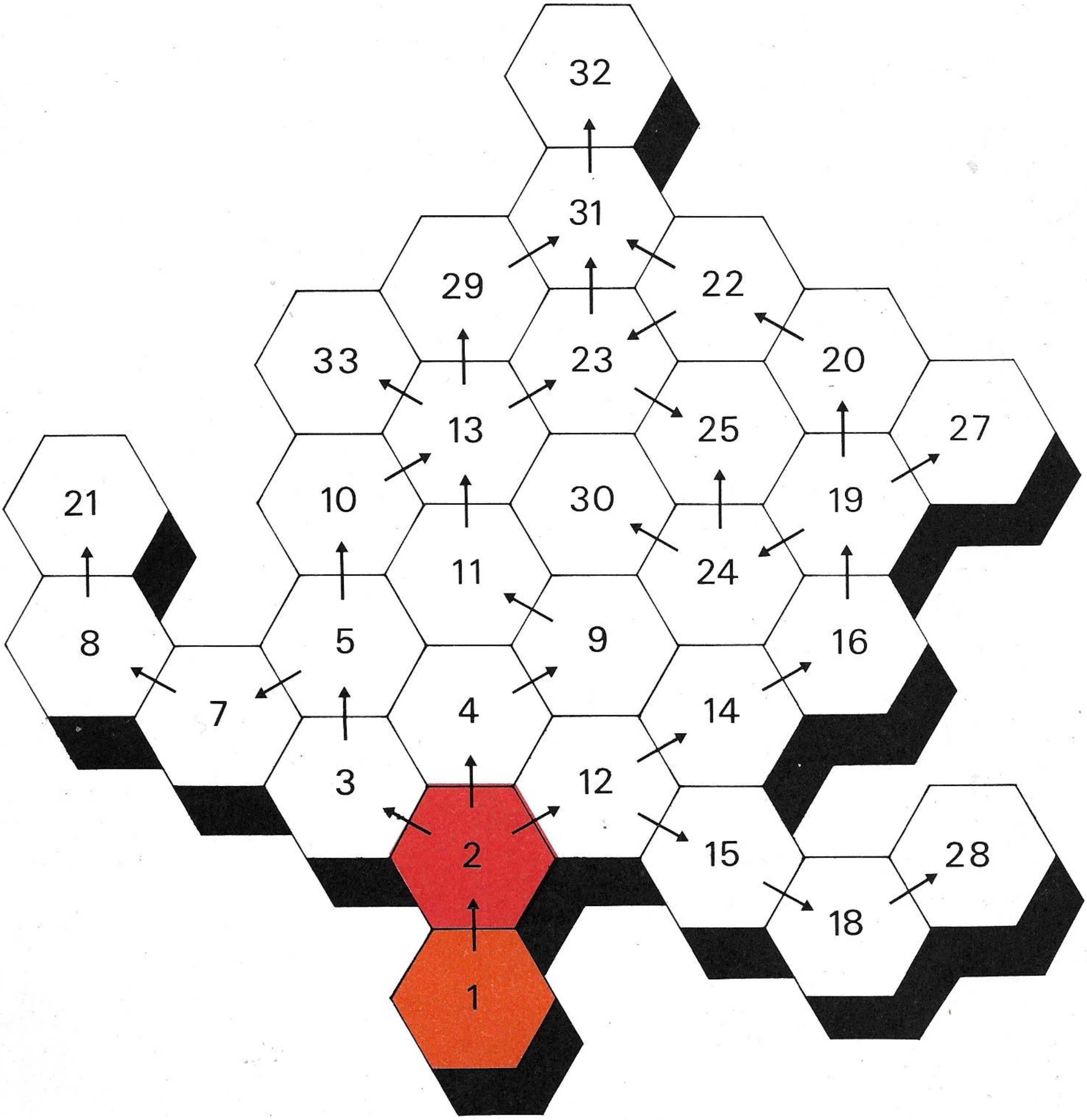
THE OPEN UNIVERSITY

Mathematics: A Second Level Course

Linear Mathematics Units 1 and 2

# Vector Spaces

# Linear Transformations



The Open University

*Mathematics: A Second Level Course*

*Linear Mathematics Unit 1*

# VECTOR SPACES

*Prepared by the Course Team*

The Open University Press

The Open University Press Walton Hall Milton Keynes MK7 6AA

First published 1971. Reprinted 1975, 1976


Designed by the Media Development Group of the Open University.

Printed in Great Britain by
Martin Cadbury

SBN 335 01090 3

This text forms part of the correspondence element of an Open University Second Level Course. The complete list of units in the course is given at the end of this text.

For general availability of supporting material referred to in this text, please write to the Director of Marketing, The Open University, P.O. Box 81, Walton Hall, Milton Keynes, MK7 6AT.

Further information on Open University courses may be obtained from the Admissions Office, The Open University, P.O. Box 48, Walton Hall, Milton Keynes, MK7 6AB.

1.3

# Contents

## Set Books

D. L. Kreider, R. G. Kuller, D. R. Ostberg and F. W. Perkins, *An Introduction to Linear Analysis* (Addison-Wesley, 1966).
E. D. Nering, *Linear Algebra and Matrix Theory* (John Wiley, 1970).

It is essential to have these books; the course is based on them and will not make sense without them.

## Conventions

Before working through this correspondence text make sure you have read *A Guide to the Linear Mathematics Course*. Of the typographical conventions given in the Guide the following are the most important.

The set books are referred to as:

K for *An Introduction to Linear Analysis*
N for *Linear Algebra and Matrix Theory*

All starred items in the summaries are examinable.

References to the Open University Mathematics Foundation Course Units (The Open University Press, 1971) take the form *Unit M100 3, Operations and Morphisms.*

## 1.0 INTRODUCTION

This first unit sets the stage for the remainder of the course. In a sense, it is a review of the basic ideas connected with a vector space which were included or hinted at in the Foundation Course, M100, in particular in *Units M100 22* and *M100 23*, *Linear Algebra I* and *II*. However, there is an important difference here in that we are now going to be more precise in our definitions and in the reasoning based on them. We shall be applying the theory of vector spaces to a number of very different examples, and it is important to be quite clear to just what extent we can draw on the theory of vector spaces in each of these examples. This can only be done by developing the theory in a rigorous way, so that we know just which theorems are provable for *all* vector spaces and which theorems only apply to vector spaces with a particular property. We therefore begin by giving a precise specification, by means of a system of axioms, of what we mean by a vector space, and we introduce some of the most important examples of vector spaces.

After this we develop the theory by looking at some of the elementary consequences of the axioms. One of the first significant results is a theorem which makes it meaningful to speak of the number of dimensions of a vector space. Just like the 3-dimensional space we live in, a vector space has a definite number of dimensions (indeed, this is one of the reasons for calling it a "space"). The number of dimensions need not be 3, however; it can be any positive integer, and there are also "infinite-dimensional" vector spaces, which we shall study later in the course.

Finally we shall look at the subspaces of a vector space: these are simply subsets of a vector space which themselves satisfy the vector space axioms. Again, it is possible to prove some useful theorems about subspaces, but in this unit we shall not go deeply into the theory of subspaces. The material in this unit covers most of Chapter I of both **K** and **N**; but we shall not be asking you to read every word in these chapters. (If you find the unit easy and have the time to spare, you might like to read some of those parts of **K** and **N** which we have not asked you to read.) Both books include exercises and some of our exercises are different from these; so you will have an ample supply of problems to work through if you feel you need more practice with any point.

## 1.1 VECTOR SPACES

### 1.1.0 Introduction

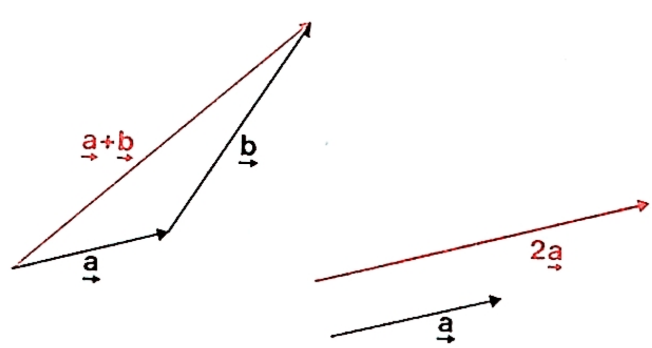
The concept of a vector space is the basis of the entire Linear Mathematics Course. It will be used, explicitly or implicitly, in every one of the units of the course. We can get a rough idea of a vector space if we regard it as a set whose elements can be (a) added together, and (b) multiplied by a real number, such that the result in either case is always an element of the set. For example, the set of ordered pairs of real numbers can be converted into a vector space by defining addition and multiplication by a real number as in the following examples:

$$(1, 2) + (\tfrac{1}{3}, \tfrac{1}{2}) = (1\tfrac{1}{3}, 2\tfrac{1}{2})$$

$$\pi \times \left(\frac{1}{3}, \frac{1}{2}\right) = \left(\frac{\pi}{3}, \frac{\pi}{2}\right)$$

If you took the Foundation Course, M100, you will already have met several other examples of vector spaces. In fact, vector spaces are introduced there (in *Unit M100 22, Linear Algebra I*) by considering a particular example, a space of geometric vectors. (A geometric vector is defined as a set of arrows all having the same length and direction. If you need to revise geometric vectors, read Section 1–8 of **K**.)

The approach to the subject used in the Foundation Course was to define geometric vectors and a way of adding them and a way of multiplying them by real numbers.




This addition and multiplication were then shown to satisfy certain rules, and in Section *22.2.2* of *M100 22* these rules were listed and called axioms of a vector space. Since we are now trying to be more careful with our mathematical statements, we shall work the other way round and shall say that a vector space is *defined* as a set with addition and multiplication which satisfy these axioms. Then whatever we deduce from the axioms must hold (by definition) in *every* vector space. In the coming sections we first introduce these axioms and then consider some examples of vector spaces.

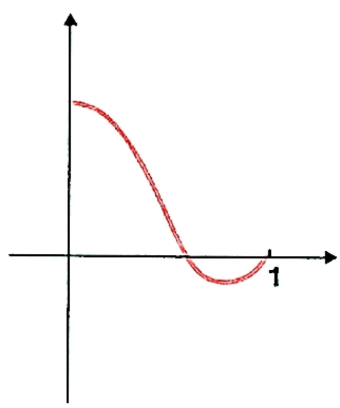
## 1.1.1 The Axioms of a Real Vector Space

A good introduction to the axioms which define a vector space is given in K. The authors' approach is similar to that of the Foundation Course.

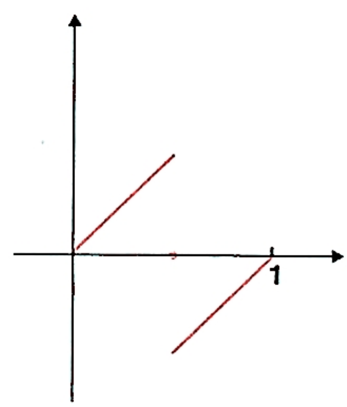
*READ Sections 1–1 and 1–2 of* K, *pages* K1–8.*

### Notes

(i) *line 1, page* K1 $R^2$ was often referred to as $R \times R$ in the Foundation Course. Remember that $R$ denotes the set of all real numbers. Whereas K uses a script type-face $\mathscr{R}$, we used an ordinary italic $R$. Also K uses bold-face type for vectors, whereas in the Foundation Course we underlined our letters to denote vectors.
(ii) *line 8, page* K3 What is meant here is that $\mathcal{C}[a, b]$ consists of all those functions whose domain is $R$ or a subset of $R$ which includes the subset $[a, b] = \{x : a \leqslant x \leqslant b, x \in R\}$ and whose codomain is $R$, which are continuous at every point of $[a, b]$. Continuity can be thought of as meaning that the graph of each function in $\mathcal{C}[a, b]$ does not have any breaks in it. $\mathcal{C}[a, b]$ will be mentioned again in the first unit on differential equations.




Graph of function $\in C[0,1]$




Graph of function $\notin C[0,1]$

(iii) *line* $-11$, *page* K3 (*i.e. 11 lines from the bottom of page 3*) What is meant here is that for $\mathcal{C}[a, b]$ to be a vector space, $\mathbf{f} + \mathbf{g}$ must also lie in the space; i.e. it must be a continuous function with domain $[a, b]$ and codomain $R$. This property of $\mathbf{f} + \mathbf{g}$ is intuitively obvious from the graphs on page K3, but to *prove* it we need the theorem that if $\mathbf{f}$ and $\mathbf{g}$ are continuous then so is $\mathbf{f} + \mathbf{g}$. This is a theorem in calculus which will not be covered in this course.
(iv) *Definition 1–1 on page* K6–7 Notice that there are two "hidden" axioms in the statements preceding axiom (i) and axiom (v). The two hidden axioms, which are really the most important of all, assert first that a vector space is closed under addition and secondly that it is closed under scalar multiplication. This accounts for the fact that in the Foundation Course we listed ten axioms (see *Unit M100 22*), whereas here we have apparently only eight.
(v) *line 1, page* K7 The real numbers are called the *scalars* for the vector space $\mathscr{V}$ so that this *scalar product* is really multiplication of a vector by a scalar.

The main point in the passage you have just read is **Definition 1–1** on page K6. This definition incorporates all that we shall mean when we use the phrase *real vector space.* If this definition seems rather arid when read in isolation, try thinking of some examples of a vector space when reading through the axioms. (You will find a selection of such examples in section *22.2.1* of *Unit M100 22.*) This will be particularly helpful in studying theorems about vector spaces.

There are three types of vector space used as examples in this passage, and we would like to bring them to your attention again because of their importance throughout the course. The first of these types comprises the spaces $R^n$ introduced in Example 2 on page K7. For any positive integer $n$,

* Remember that we do not intend you to work the exercises in a reading section unless specifically mentioned. Also remember that you were advised to mark the beginning, end and lines where notes occur before starting to read.

the elements of $R^n$ are $n$-tuples, that is, arrays of $n$ real numbers $(x_1, x_2, \ldots, x_n)$. If $n = 1, 2, 3$ we have alternative visual interpretations of $R^n$. If $n = 1$, an element of $R^1$ is a " 1-tuple," $(x_1)$, and if we ignore the brackets we see that $R^1$ can be thought of as being the same as $R$, which is a vector space (as was pointed out in Example 1 on page K7). $R$ can be visualized as the set of points on a line. The elements of $R^2$ are ordered pairs, which can be visualized as points in a plane; this was discussed in Section 1–1 of K. The elements of $R^3$ are ordered triples of real numbers and $R^3$ can be visualized as points in three-dimensional space. These last three examples are particularly useful as they give a way of visualizing the properties of vector spaces.

The next type comprises spaces of the form $C[a, b]$ whose elements are functions. We shall be particularly interested in this vector space in the units on differential equations in this course.

The third type is the vector space $P$ introduced in Example 3 on page K8. An element of $P$ is a polynomial, that is, an expression of the form

$$a_0 + a_1x + \cdots + a_nx^n$$

where $x$ is a variable (called an *indeterminate*), $n$ is a positive integer or zero which, if $a_n \neq 0$, is called the *degree* of the polynomial, and $a_0, a_1, \ldots, a_n$ are real numbers called the *coefficients*. The addition of polynomials and their multiplication by real numbers is defined in the usual way; for example

$$(1 + 2x) + (3x + x^2) = 1 + 5x + x^2$$
$$(-1) \times (3\tfrac{1}{2} - 2x^3) = -3\tfrac{1}{2} + 2x^3.$$

Note that if all the coefficients are zero we obtain the *zero polynomial*, 0. The zero polynomial is a member of $P$, and in future when a set of polynomials such as $P$ is referred to as a vector space, the zero polynomial is always included.

Every polynomial is associated with a function of the form

$$p: x \longmapsto a_0 + a_1x + \cdots + a_nx^n \qquad (x \in R)$$

which is called a *polynomial function*. We often use the same symbol, $p$ say, to stand for the polynomial and the associated polynomial function.

### Exercises

Test your understanding of what you have read by working the following exercises. (We suggest the minimum number of exercises that we consider you should work. If you feel you need further practice, then try some more.) The solutions are printed immediately after the exercises.

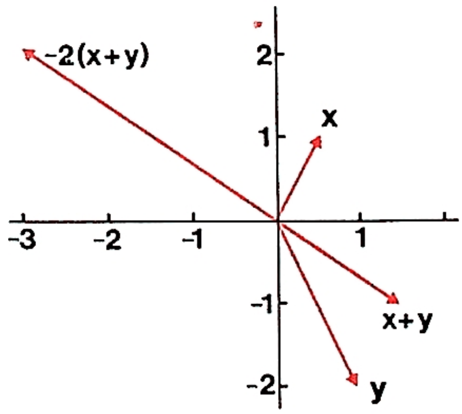
1. (i) Page K5, Exercise 2.
   (ii) Page K5, Exercise 6.
2. (i) Page K5, Exercise 12.
   (ii) Page K6, Exercise 19.
   (Note that K does not always distinguish between a function and its image. Thus $\tan x$ is used both as a function and the image of $x$ under the function tan. Also the domain is not explicitly stated: it is assumed that the domain is as large as possible. Thus the domain of $\tan x$ is $\{x : x \in R, \cos x \neq 0\}$.)
3. (i) Page K9, Exercise 4(a).
   (ii) Page K9, Exercise 5(a).
4. Page K9, Exercise 6.

**Solutions**

The answers to odd-numbered exercises in K can be found at the back of K, page 725 onwards. We shall add notes to these answers where we think they might be helpful. We shall also give answers to those even-numbered exercises which we ask you to work.

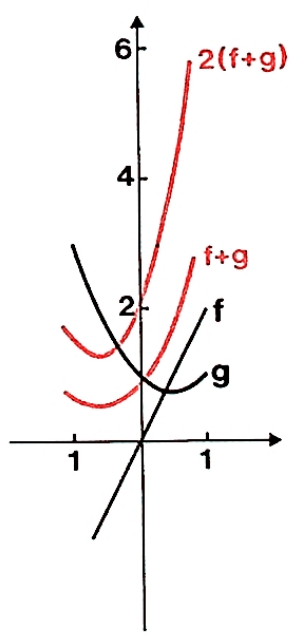
1. (i) $\mathbf{x} + \mathbf{y} = (1\frac{1}{2}, -1)$

$\alpha(\mathbf{x} + \mathbf{y}) = (-3, 2)$

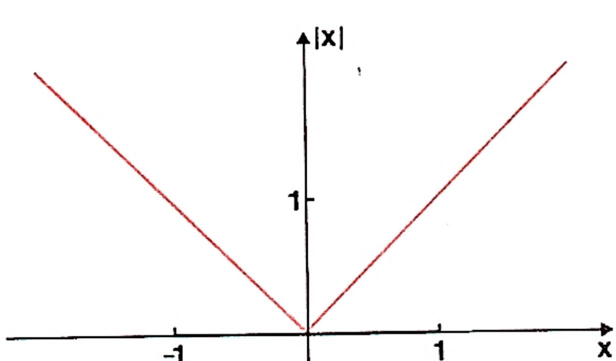



(ii) $(\mathbf{f} + \mathbf{g})(x) = x^2 + x + 1$

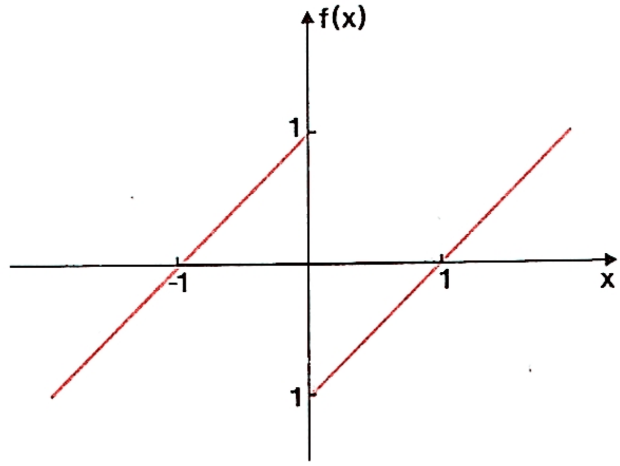
$\alpha(\mathbf{f} + \mathbf{g})(x) = 2x^2 + 2x + 2$




2. (i) This function belongs to $\mathcal{C}[-1, 1]$.



(ii) The function does not belong to $\mathcal{C}[-1, 1]$ because it is discontinuous at 0.




3. (i) $\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \alpha_3 \mathbf{x}_3 = (0, -1, 4, -2)$

   (ii) $\alpha_1 \mathbf{p}_1 + \alpha_2 \mathbf{p}_2 + \alpha_3 \mathbf{p}_3 = -x^3 + 2x^2 - 8x + 3$

4. This is not a vector space because multiplying such a polynomial by a non-integral scalar may give a polynomial with non-integral coefficients; e.g. if

$$\alpha = \tfrac{1}{2} \quad \text{and} \quad \mathbf{p} = x^2 + 2x + 3,$$

then

$$\alpha\mathbf{p} = \tfrac{1}{2}x^2 + x + \tfrac{3}{2}$$

which is not in the set specified.

### 1.1.2 Fields

We have so far seen the definition of a *real* vector space, but this is not the most general type of vector space. In a real vector space, we require that if $\mathbf{x}$ is in the space then so is $\alpha\mathbf{x}$ with $\alpha$ any real number; but instead we could have specified a different set of possibilities for $\alpha$, for example, complex numbers or rational numbers. (Complex vector spaces, in which $\alpha$ is allowed to be any *complex* number, are very useful in solving linear differential equations, as we shall see later in the course.) To define a general vector space we replace the requirement that $\alpha$ be real by a requirement that it be drawn from a set satisfying a certain set of axioms, which are satisfied by the real numbers, the complex numbers, the rational numbers, and some other more special sets too. Any set satisfying these axioms is called a *field* and its elements (in the context of vector spaces) are called *scalars*.

To see the details of how a field is defined, we turn to N.

*READ from the word* **Definition** *on page* N6 *as far as* **Definition** *on page* N7.

The 11 axioms for a field together with the 10 axioms of a vector space may seem a formidable list. In a little while you will not find this so. It becomes a habit to think in terms of specific examples, even when carefully proving very general theorems. Also, the notation helps us remember the rules; we use the ordinary notation for multiplication and addition.

You may find it helpful to notice that the first five field axioms $F1$–$5$ in N are merely stating that $(F, +)$ is a commutative group. (See *Unit M100 30*: the concept of a group is not essential to this course, although it can sometimes be useful. See also page N8.) The next five, $F6$–$10$, state that if we remove 0 from $F$, then the remainder is a commutative group for multiplication. So we can replace the 11 axioms by 3, viz:

(i) $(F, +)$ is a commutative group
(ii) $(F_1, \cdot)$ is a commutative group, where $F_1$ is $F$ without the identity 0 for $+$
(iii) $\cdot$ is distributive over $+$.

We do not ask you to remember the set of axioms of a field; we merely wish to draw your attention to its existence and the fact that any results proved for vector spaces in general apply to vector spaces over any field. In most practical cases the field is the set of real numbers or the set of complex numbers.

## 1.1.3 The Axioms for a General Vector Space

We next have the axioms for a general vector space; that is, a vector space with a general field $F$ as the scalars.

*READ from* **Definition** *on page* N7 *to "... satisfying these conditions" about halfway down page* N8.

If we compare the axioms listed on pages K6–7 with those listed in N then there are three points to notice. The most obvious is in the notation: K uses bold-face roman letters for vectors and Greek letters for scalars but N uses Greek for vectors and italic for scalars. This is just a notational difference and nothing more. The next point is that N has two more axioms than K, namely $A1$ and $B1$. These axioms (which we have already referred to earlier in a note on K) are sometimes called *closure* axioms and make explicit the requirement that addition and scalar multiplication do give elements of the vector space. Finally, there is the difference we have already mentioned that K is defining a *real* vector space and therefore requires $F$ to be the field of real numbers, whereas N defines a general vector space and so $F$ is any field.

## 1.1.4 Examples of Vector Spaces

The next reading passage lists nine examples of vector spaces.

*READ the rest of page* N8 *and continue to the end of Example (9) on page* N9.

Do not worry about the details of Examples (4), (5), (6), (7) at this stage but pay particular attention to Examples (1), (3), (8), (9). What do you notice about these latter examples?

### Notes

(i) *line −12, page* N8 For the terminology relating to polynomials, see the end of sub-section 1.1.1 of this text.
(ii) *line −4, page* N8 A "real-valued function of a real variable" is what we called a *real function* in M100. That is, a function whose domain and codomain are $R$ or subsets of $R$. The "values" of a function are what we called its images in M100.

The thing you should notice about Examples (1), (3), (8) and (9) is that they are the same as the examples we met in K but with a general field $F$ of scalars instead of the field $R$ of real numbers used in K. Example (2) is an obvious extension of Example (1). Whatever $n$ we use, all the vectors in $P_n$ are also in $P$; for example, $P_2$ consists of all polynomials in $x$ of the form $a + bx$, where $a$, $b$ are in the field $F$. Examples (4), (5), (6) require theorems to verify $A1$ and $B1$ and these may be theorems you have not yet met (and which will not be proved in this course). Example (4), for example, requires the theorem referred to earlier, that the sum of two continuous functions is itself continuous. Example (7) is particularly interesting because it illustrates the possibility of applying vector space theory to the study of differential equations. This application was found very fruitful in *Unit M100 31, Differential Equations II*, and we shall exploit it further in the units on differential equations in this course.

### Exercise

With the notation as in N, let $V$ be the set of all polynomials of degree 3 and let $F = R$. Is $V$ a real vector space? (*Hint:* Compare this set with Example 2 on page N8.)

**Solution**

There are a number of reasons why this is not a vector space. For instance, $A1$ is not satisfied: for example, if

$$\alpha = x^3 + 3x^2 + 2$$

and

$$\beta = -x^3 + 3x + 5,$$

then $\alpha + \beta$ is a polynomial of degree 2 and so does not belong to $V$. Another reason is that $A3$ is not satisfied.

### 1.1.5 The Zero Vector and the Negative of a Vector

The last passage in the current section of N covers some simple manipulations based on the axioms defining a vector space. You may like to compare it with Section 1–3 of K. You may find the one-line proofs hard to follow, and there is no need to follow the argument in detail; we do ask you, however, to note the results which we summarize below. We shall be commenting on one or two such proofs later on.

*READ the remainder of Section* 1 *on pages* N9–10.

The results proved in this section are:

(i) the zero vector is unique;
(ii) the negative (additive inverse) of a vector is unique;
(iii) $0\alpha = 0$, for any vector $\alpha$ (the zero scalar times any vector is equal to the zero vector);
(iv) $a0 = 0$, for any scalar $a$ (any scalar multiple of the zero vector is equal to the zero vector);
(v) $(-1)\alpha = -\alpha$ (see $F8$ and $F4$ for the definition of $-1$).

### 1.1.6 Summary of Section 1.1

This section contains the axioms for a *field* and the axioms for a *vector space*.

A *field* is any mathematical structure satisfying the axioms $F1$ to $F11$ on page N6–7. The only fields of importance in this course are the real numbers and the complex numbers. * * *

A *vector space* (over some given field $F$) is any mathematical structure satisfying the axioms $A1$ to $A5$ and $B1$ to $B5$ on pages N7–8. In particular, the set is closed under the operations of addition, and multiplication by any scalar (i.e. element of $F$). * * *

You are not expected to memorize these lists of axioms, but you should be able to use them to check whether a given structure is, or is not, a vector space.

This section also introduces three important examples of (real) vector spaces:

$R^n$, the space of $n$-tuples $(a_1, a_2, \ldots, a_n)$ of real numbers * * *
$C[a, b]$, the space of functions continuous on the interval $[a, b]$ * * *
$P$, the space of polynomials with real coefficients * * *

## 1.2 LINEAR DEPENDENCE AND INDEPENDENCE

### 1.2.0 Introduction

Having defined a vector space as (loosely) a set of objects which is closed under the operations of addition, and multiplication by scalars, we now look at some of the consequences of our definition. In particular, if $\alpha$ and $\beta$ are two given vectors, what can we form from them using the two vector space operations, one after the other? For example, we could add $\alpha$ to $\beta$ and multiply the result by a scalar, obtaining a vector such as

$$2(\alpha + \beta) = 2\alpha + 2\beta;$$

or we could multiply both $\alpha$ and $\beta$ by scalars and then add the resulting vectors, obtaining a vector such as

$$3\alpha + (-\tfrac{1}{2})\beta = 3\alpha - \tfrac{1}{2}\beta.$$

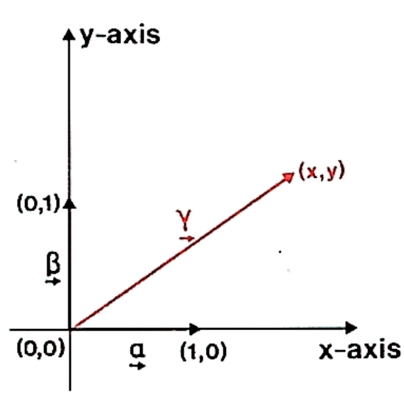
Expressions of the form $2\alpha + 2\beta$ or $3\alpha - \frac{1}{2}\beta$, or in general $a\alpha + b\beta$, where $a$ and $b$ are scalars, are called *linear combinations* of the two vectors $\alpha$ and $\beta$. Similarly, the most general linear combination of three given vectors $\alpha$, $\beta$ and $\gamma$ is

$$a\alpha + b\beta + c\gamma,$$

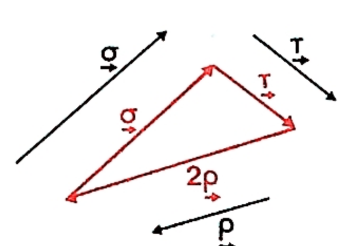
where $a$, $b$ and $c$ are scalars. This idea of taking linear combinations of elements of a given set of vectors is very useful because it combines the two basic operations of addition and scalar multiplication; for example, we can now define a vector space (loosely) as a set of objects which is closed under the "operation" of taking linear combinations.

Given any set of vectors in a vector space, we can ask the following two questions about it.

1. Can every non-zero vector in the space be expressed as a linear combination of the vectors in the set? If the answer is yes, the set is said to *span* the space.
2. Can the zero vector be expressed as a linear combination of the vectors in the set, in a non-trivial way (that is, in a way other than $0 = 0\alpha + 0\beta + \cdots + 0\gamma$)? If the answer is yes, the vectors are said to be *linearly dependent*.




The geometric vectors $\underset{\rightarrow}{\alpha}$ and $\underset{\rightarrow}{\beta}$ span the plane of the paper, because any other geometric vector $\underset{\rightarrow}{\gamma}$ in the plane can be written in the form $\underset{\rightarrow}{\gamma} = x\underset{\rightarrow}{\alpha} + y\underset{\rightarrow}{\beta}$




The geometric vectors $\underset{\rightarrow}{\sigma}$, $\underset{\rightarrow}{\rho}$, and $\underset{\rightarrow}{\tau}$ are linearly dependent, because $\underset{\rightarrow}{\sigma} + \underset{\rightarrow}{\tau} + 2\underset{\rightarrow}{\rho} = \underset{\rightarrow}{0}$

Intuitively, it would seem that the more vectors there are in the set, the more linear combinations it has and therefore the more likely it is to span

the space and the more likely it is to be linearly dependent; thus we might expect to find some connection between the number of vectors in a *spanning set* (i.e. a set of vectors which spans the space) and the number in a linearly dependent set. Such a connection does exist, and the present section leads up to a theorem which provides it. It is called the *Steinitz Replacement Theorem* and is the first important theorem in this course.

This section is based on Section 2 of Chapter I of **N**. The same material is also treated on pages 18–20 of **K**. We shall not specifically ask you to read the treatment in **K**, but if you are having difficulty you may find it helpful to consult **K** for a different point of view.

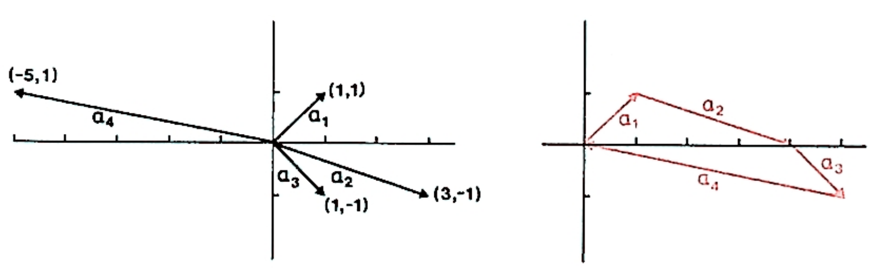
If you have had a quick look at Section 2 (page N10 onwards), you will have seen that there are seven theorems, each with its proof. In fact, six of these proofs only deduce simple facts about linear dependence and span, some of which are used in the proof of the important theorem of the section, ***Theorem 2.7.*** We shall not ask you to read all these theorems and their proofs. What we shall do in this section is to illustrate the ideas we will need to prove the key theorem of this section, the Replacement Theorem.

## 1.2.1 Linear Dependence

In this sub-section we are interested in linear combinations of vectors which give us the zero vector. We first look at an example.

Let us look at the real vector space $R^2$: the vectors in $R^2$ are ordered pairs of real numbers, $(a, b)$. We pick four vectors from $R^2$: $\alpha_1 = (1, 1)$, $\alpha_2 = (3, -1)$, $\alpha_3 = (1, -1)$ and $\alpha_4 = (-5, 1)$. We have chosen them carefully: if we add them we get

$$\begin{aligned}\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 &= (1, 1) + (3, -1) + (1, -1) + (-5, 1)\\ &= (1 + 3 + 1 - 5, 1 - 1 - 1 + 1)\\ &= (0, 0)\\ &= 0, \text{ the zero vector in } R^2.\end{aligned}$$




Thus there is a non-trivial linear combination of $\alpha_1, \ldots, \alpha_4$ which is equal to the zero vector. In other words, $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are *linearly dependent*. We could write this in another way:

$$\alpha_4 = -(\alpha_1 + \alpha_2 + \alpha_3),$$

which tells us that $\alpha_4$ is a linear combination of $\alpha_1, \alpha_2, \alpha_3$. This is no accident: if a set of vectors is linearly dependent, then it is always possible to express one of them as a linear combination of the others. (We illustrate that not every vector in a linearly dependent set can be expressed in terms of the others by the following linearly dependent set in $R^3$

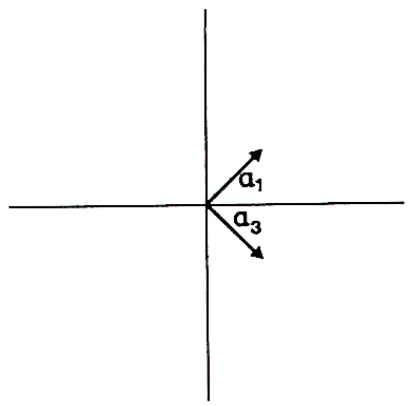
$$\{(1, -1, 0), (-2, 2, 0), (1, 1, 1)\}.$$

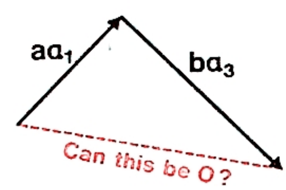
We have linear dependence since

$$2(1, -1, 0) + (-2, 2, 0) + 0(1, 1, 1) = (0, 0, 0).$$

But it is not possible to express $(1, 1, 1)$ as a linear combination of the others.)

Suppose now that we consider a smaller set of vectors, say $\alpha_1$ and $\alpha_3$, and try to find a *linear relation* between them—that is, a linear combination that is equal to the zero vector.




This means looking for real numbers $a$ and $b$ such that

$$a\alpha_1 + b\alpha_3 = 0$$

i.e.

$$a(1, 1) + b(1, -1) = (0, 0).$$

Carrying out the indicated scalar multiplications and additions on the left side, we get

$$\begin{aligned} a(1, 1) + b(1, -1) &= (a, a) + (b, -b) \\ &= (a + b, a - b). \end{aligned}$$

If there is a linear relation between $\alpha_1$ and $\alpha_3$, this vector $(a + b, a - b)$ must be the vector $(0, 0)$. This implies that the corresponding components are the same, i.e. that

$$a + b = 0$$

and

$$a - b = 0,$$

from which it follows that $a = 0$ and $b = 0$. Thus the only linear relation between $\alpha_1$ and $\alpha_3$ is the trivial one,

$$0\alpha_1 + 0\alpha_3 = 0.$$

The set of vectors $\{\alpha_1, \alpha_3\}$ is therefore *not* linearly dependent; we say that it is *linearly independent*.

The method of generalizing these ideas to arbitrary sets of vectors in an arbitrary vector space is described in the next passage of N.

*READ Section* 2 *pages* N10–12, *up to but not including Theorem 2.1, omitting the paragraph beginning "It is clear that...." on page* N11.

**Notes**

(i) *line 1, page* N11 We require that only a finite number of coefficients in an expression $\sum_i a_i \alpha_i$ are non-zero, because we have as yet no way to give any meaning to the sum of an infinite set of vectors. We shall see how to deal with infinite sums later in the course.

(ii) *line 17, page* N11 This piece of proof has been printed on one line to save space. It may help you if you write such proofs out for yourself on several lines. For example, this particular line could be rewritten to distinguish each step in the following way:

"For if $a\alpha = 0$ with $a \neq 0$, then

| | |
|---|---|
| $\alpha = 1 \cdot \alpha$ | (by axiom *B5*) |
| $= (a^{-1} \cdot a)\alpha$ | (by *F9*, since $a \neq 0$) |
| $= a^{-1}(a\alpha)$ | (by *B2*) |
| $= a^{-1}(0)$ | (since $a\alpha = 0$) |
| $= 0.$" | (by the first paragraph on page N10) |

A point to note about this proof is that it is an indirect proof. (See *Unit M100 17*.) To see why it is a proof it is perhaps simplest to convert it to a proof by contradiction.

We start with the hypothesis $a \neq 0$, and by N's argument reach $\alpha = 0$. This contradicts the original assertion that our set consists of a non-zero vector, $\alpha$. Thus our original hypothesis, $a \neq 0$, is disproved; i.e. we have proved $a = 0$.

Alternatively, do not read such proofs at all but try to prove the result oneself. This will often require more time and may sometimes lead to a different proof altogether. For instance, in this case, one might well come up with the following proof.

| | |
|---|---|
| $a\alpha = 0$ | $a \neq 0$ |
| $\Rightarrow a^{-1}(a\alpha) = a^{-1}0$ | (*F9*) |
| $\Rightarrow (a^{-1}a)\alpha = 0$ | (*B2* and first paragraph on page N10) |
| $1\alpha = 0$ | (*F9*) |
| $\alpha = 0$ | (*B5*) |

(iii) *line 2, page* N12 $\{\alpha_i\}$ is a notation for a set of vectors. The letter $i$ is allowed to range over the integers. Thus both the following sets could be denoted by $\{\alpha_i\}$

$\{\alpha_i\} = \{\alpha_1, \alpha_3, \alpha_5\}$, a finite set

$\{\alpha_i\} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \ldots\}$, an infinite set.

(iv) *line 10, page* N12 Example (1) is on page N8. A *monomial* is a polynomial whose coefficients are all zero except one.

### Exercise

Page N14, Exercise 1.

### Solution

The answers to selected exercises can be found at the back of N (page 325 onwards). We shall add notes to these answers where we think they might be helpful. We shall also give answers to those exercises which we ask you to work and to which N does not give answers.

For this exercise, the answer given by N is brief. Here is a fuller solution.

Can we find real numbers $a$, $b$, $c$ and $d$, not all zero, such that

$$a(x^2 + x + 1) + b(x^2 - x - 2) + c(x^2 + x - 1) + d(x - 1)$$

is the zero polynomial?

Rearranging gives

$$(a + b + c)x^2 + (a - b + c + d)x + (a - 2b - c - d)$$

This expression is the zero polynomial if, simultaneously,

$$\begin{aligned} a + b + c &= 0 \\ a - b + c + d &= 0 \\ a - 2b - c - d &= 0. \end{aligned}$$

Since there are 4 unknowns and only 3 equations, there is not enough information to determine $a$, $b$, $c$, $d$ uniquely. If we choose a value for one of them arbitrarily, the other values are then uniquely determined.

Let $d = 1$, say.

Hence we obtain $a = \frac{3}{4}, b = \frac{1}{2}, c = -1\frac{1}{4}$.

Thus the given polynomials are linearly dependent, and we can express $p_4$, say, as a linear combination of the other three:

$$p_4 = -\tfrac{3}{4}p_1 - \tfrac{1}{2}p_2 + 1\tfrac{1}{4}p_3.$$

### 1.2.2 A Theorem on Linear Dependence

The object of this sub-section is to understand ***Theorem 2.1*** on page N12.

*READ Theorem 2.1 and its proof on page* N12,

but if you are not sure that you have understood the statement of the theorem or its proof, work the following exercises and then read the theorem again.

**Exercises**

1. For the following vectors in $R^2$

$$\beta_1 = (1, 1), \qquad \beta_2 = (3, -1), \qquad \beta_3 = (1, -1)$$

and $\alpha = (6, -4)$, we can easily show that

$$\alpha = -\beta_1 + 2\beta_2 + \beta_3$$

i.e. $\alpha$ is a linear combination of $\{\beta_i\}$.

If further

$$\gamma_1 = (1, 0), \qquad \gamma_2 = (0, 1)$$

express each $\beta$ as a linear combination of $\{\gamma_j\}$ and hence express $\alpha$ as a linear combination of the $\{\gamma_j\}$.

(Although this is not the quickest way to express $\alpha$ in terms of the $\{\gamma_j\}$, we want to illustrate the theorem.)

2. Let $\gamma_1 = 1 + x$, $\gamma_2 = 1 - x^2$, $\gamma_3 = 1 + x - x^2$ be three vectors in the real vector space $P_3$ of polynomials in $x$ of degree at most 2.

   (i) Write the vectors $\beta_1 = 1, \beta_2 = x, \beta_3 = x^2$ as linear combinations of $\{\gamma_1, \gamma_2, \gamma_3\}$. (*Hint*: try adding $\gamma_1$ and $\gamma_2$.)

   (ii) Use your solution to (i) to write

$$\alpha = 2 + 3x - x^2$$

   as a linear combination of $\{\gamma_1, \gamma_2, \gamma_3\}$.

**Solutions**

1. $$\beta_1 = \gamma_1 + \gamma_2, \beta_2 = 3\gamma_1 - \gamma_2, \beta_3 = \gamma_1 - \gamma_2.$$

   Therefore, since

$$\begin{aligned} \alpha &= -\beta_1 + 2\beta_2 + \beta_3, \\ \alpha &= -(\gamma_1 + \gamma_2) + 2(3\gamma_1 - \gamma_2) + (\gamma_1 - \gamma_2) \\ &= -\gamma_1 - \gamma_2 + 6\gamma_1 - 2\gamma_2 + \gamma_1 - \gamma_2 \\ &= 6\gamma_1 - 4\gamma_2. \end{aligned}$$

   If you are using this exercise to understand the theorem, then you should notice that $\alpha = \sum_i b_i \beta_i$,

   where

$$b_1 = -1, b_2 = 2, b_3 = 1, \text{ and } \beta_i = \textstyle\sum_j c_{ij}\gamma_j,$$

   where

$$c_{11} = c_{12} = 1;\ c_{21} = 3,\ c_{22} = -1,\ c_{31} = 1,\ c_{32} = -1.$$

2. (i) $$\gamma_1 + \gamma_2 = 2 + x - x^2$$
$$= 1 + \gamma_3.$$

Hence $\beta_1 = \gamma_1 + \gamma_2 - \gamma_3$.

$$\gamma_1 = 1 + x$$
$$= \beta_1 + \beta_2$$
$$= \gamma_1 + \gamma_2 - \gamma_3 + \beta_2$$

so that $\beta_2 = \gamma_3 - \gamma_2$.

Finally

$$\beta_3 = -\gamma_2 + 1$$
$$= -\gamma_2 + \gamma_1 + \gamma_2 - \gamma_3$$
$$= \gamma_1 - \gamma_3.$$

That is $\beta_1 = \gamma_1 + \gamma_2 - \gamma_3$, $\beta_2 = \gamma_3 - \gamma_2$, $\beta_3 = \gamma_1 - \gamma_3$

(ii) $$\alpha = 2 + 3x - x^2$$
$$= 2\beta_1 + 3\beta_2 - \beta_3$$
$$= 2\gamma_1 + 2\gamma_2 - 2\gamma_3 + 3\gamma_3 - 3\gamma_2 - \gamma_1 + \gamma_3$$
$$= \gamma_1 - \gamma_2 + 2\gamma_3.$$

### 1.2.3 The Set Spanned by a Set of Vectors

We have so far considered what it means to say that one vector is a linear combination of certain other vectors. What we want to consider now is what happens when we start with a set of vectors, $A$ say, and form all the possible linear combinations of these vectors. We will get a larger set of vectors; we denote it by $\langle A \rangle$ and call it the set *spanned* by $A$.

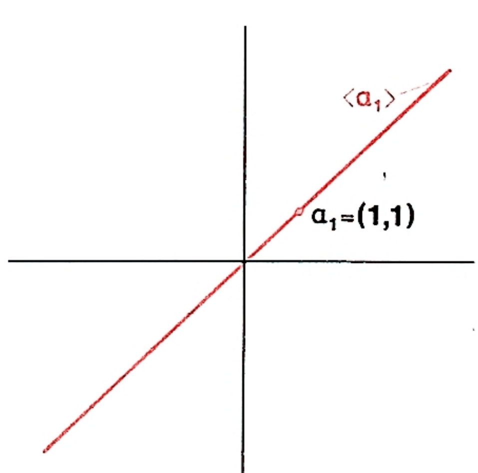
Let us illustrate this with the four vectors

$$\alpha_1 = (1, 1), \qquad \alpha_2 = (3, -1), \qquad \alpha_3 = (1, -1), \qquad \alpha_4 = (-5, 1)$$

in $R^2$. Suppose first that $A$ just contains $\alpha_1 = (1, 1)$ and we want to find all linear combinations of $\alpha_1$. From the definition of linear combination we see that any vector $\beta$ which is a linear combination of $\alpha_1$ will be of the form $\beta = a\alpha_1 = (a, a)$ where $a$ is a real number. Conversely, any vector $(a, a) = a\alpha_1$ is a linear combination of $\alpha_1$ and this shows that the set *spanned* by $\alpha_1$ is the set of vectors

$$\langle \alpha_1 \rangle = \{(a, a) : a \in R\}.$$

We can represent the elements of $R^2$ in many ways: in the diagram we show vectors as points, rather than as geometric or position vectors as before.



Now suppose we let $A = \{\alpha_1, \alpha_2\}$ and we try to find $\langle A \rangle$. Again we see from the definition of linear combination that

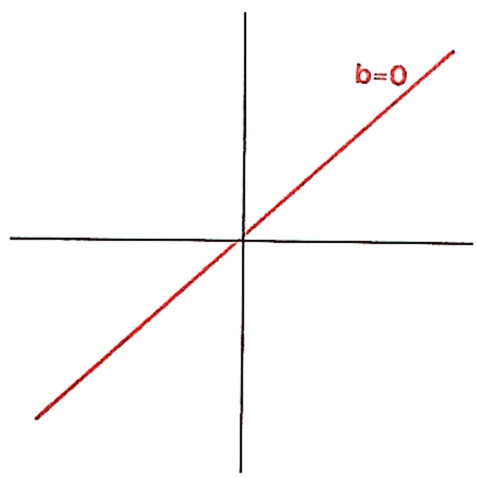
$$\langle A \rangle = \{a\alpha_1 + b\alpha_2 : a, b \in R\}.$$

That is $\langle A \rangle$ consists of *all* the vectors we can get of the form $a\alpha_1 + b\alpha_2$, for *all* possible choices of $a$ and $b$ in $R$. We look at an arbitrary vector $\beta$ in $\langle A \rangle$.

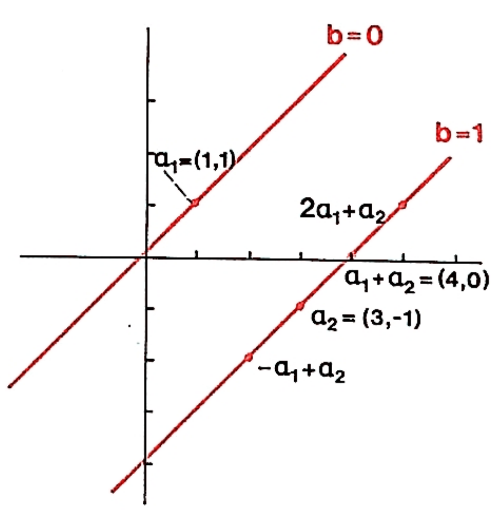
$$\begin{aligned}\beta &= a\alpha_1 + b\alpha_2 \\ &= a(1, 1) + b(3, -1) \\ &= (a + 3b, a - b).\end{aligned}$$

We therefore have $\langle A \rangle = \{(a + 3b, a - b) : a, b \in R\}$

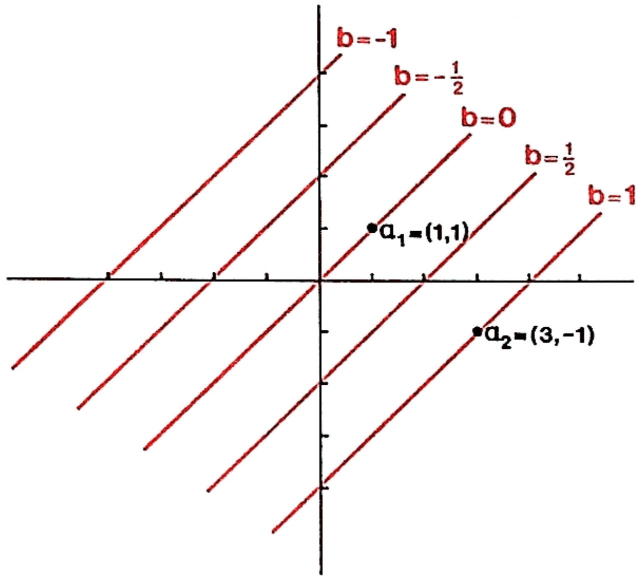
We can look at a diagram to see what is happening. When $b$ is zero we have just $\langle \alpha_1 \rangle$, which we have seen can be represented by a line in the plane.




Now suppose that $b = 1$; then we have the set of vectors $(a + 3, a - 1)$ and as we let $a$ take all values in $R$ we again get a line in our diagram.



If we try some other values for $b$, we get other lines.




It seems from the diagram that by taking all possible values for $a$ and $b$ we will get a set of lines covering the plane. That is, the diagram suggests that $\langle\alpha_1, \alpha_2\rangle^* = R^2$. Can we, in fact, prove this? In order to prove that $\langle\alpha_1, \alpha_2\rangle = R^2$ we have to show that $\langle\alpha_1, \alpha_2\rangle$ and $R^2$ contain the same vectors. Now $\alpha_1, \alpha_2$ are in $R^2$, and because $R^2$ is a vector space, any linear combination of $\alpha_1, \alpha_2$ *must* be in $R^2$. We therefore have that $\langle\alpha_1, \alpha_2\rangle \subseteq R^2$. To complete the proof that $\langle\alpha_1, \alpha_2\rangle = R^2$, we have to show that $R^2 \subseteq \langle\alpha_1, \alpha_2\rangle$; that is, that every vector in $R^2$ is in $\langle\alpha_1, \alpha_2\rangle$. Suppose that $\beta \in R^2$ and that $\beta = (c, d)$: we have to show that

$$\beta = a\alpha_1 + b\alpha_2$$

for some $a$ and $b$. That is

$$(c, d) = (a + 3b, a - b)$$

whence

$$a + 3b = c,$$
$$a - b = d.$$

Solving these simultaneous equations for $a$ and $b$ we see that $b = \dfrac{c-d}{4}$ and $a = \dfrac{c+3d}{4}$. So, if $\beta = (c, d)$, then

$$\beta = \frac{c+3d}{4}\alpha_1 + \frac{c-d}{4}\alpha_2$$

i.e. $\beta$ is in $\langle\alpha_1, \alpha_2\rangle$, and so $R^2 \subseteq \langle\alpha_1, \alpha_2\rangle$. This completes the proof that $\langle\alpha_1, \alpha_2\rangle = R^2$.

We could continue by considering $\langle\alpha_1, \alpha_2, \alpha_3\rangle$, but for the moment we have seen enough to illustrate the idea of span.

*READ the two short paragraphs on page* N12 *between* ***Theorem 2.2*** *and* ***Theorem 2.3.***

**Notes**

(i) *line* $-7$, *page* N12 Since the empty set $\emptyset$ is useful for dealing with intersections and subsets of arbitrary sets we would like to talk about $\langle\emptyset\rangle$, the set spanned by the empty set. The most sensible way to do this is to make it a definition that $\langle\emptyset\rangle$ is to be the set containing the zero vector alone.

* $\langle\alpha_1, \alpha_2\rangle$ is an abbreviated form of $\langle\{\alpha_1, \alpha_2\}\rangle$, the set of vectors spanned by the set $\{\alpha_1, \alpha_2\}$.

(ii) *line −4, page* N12 Here we see one use of symbols. The statement of ***Theorem 2.1*** takes 2 lines of print, but after defining the symbol $\langle\ \rangle$ the statement of the theorem is reduced to half a line,

"If $A \subset \langle B \rangle$ and $B \subset \langle C \rangle$ then $A \subset \langle C \rangle$."

If such lines of symbols worry you, try expressing them in words: in this case we have "If $A$ is contained in the set spanned by $B$ and $B$ is contained in the set spanned by $C$, then $A$ is contained in the set spanned by $C$." You can go one stage further and replace the words "is spanned by" to get: "If each element of $A$ is a linear combination of the elements of $B$ and if each element of $B$ is a linear combination of the elements of $C$, then ..." and we are back at the statement of ***Theorem 2.1.***

In this case our introduction and notes may seem to you to swamp the piece of text that we asked you to read so that we might have done better without N. This may or may not be true, but we remind you that one of our objectives in this course is to enable you to read ordinary mathematical texts. We hope that by the end of this course you will be able to read books like N with much less help than we are giving at present.

At the beginning of this sub-section we looked at the set spanned by $\alpha_1 = (1, 1)$ and the set spanned by $\{\alpha_1, \alpha_2\}$, $\alpha_2 = (3, -1)$, but we did not go further and look at the set spanned by $\{\alpha_1, \alpha_2, \alpha_3\}$, $\alpha_3 = (1, -1)$. Let us see what happens if we do consider $\langle(1, 1), (3, -1), (1, -1)\rangle$.

We have seen that

$$\langle\alpha_1, \alpha_2\rangle = R^2$$

and hence

$$\alpha_3 \in \langle\alpha_1, \alpha_2\rangle.$$

In fact $(1, -1) = -\frac{1}{2}(1, 1) + \frac{1}{2}(3, -1)$, i.e.

$$\alpha_3 = -\tfrac{1}{2}\alpha_1 + \tfrac{1}{2}\alpha_2 .$$

In other words, $\alpha_3$ is a linear combination of $\alpha_1$ and $\alpha_2$. A point which is perhaps not so obvious is this: suppose $\beta$ is in $\langle\alpha_1, \alpha_2, \alpha_3\rangle$; then $\beta = a\alpha_1 + b\alpha_2 + c\alpha_3$ for real numbers $a$, $b$, $c$. But $\alpha_3 = -\frac{1}{2}\alpha_1 + \frac{1}{2}\alpha_2$, and so we can substitute for $\alpha_3$.

$$\begin{aligned}\beta &= a\alpha_1 + b\alpha_2 - \frac{c}{2}\alpha_1 + \frac{c}{2}\alpha_2\\ &= \left(a - \frac{c}{2}\right)\alpha_1 + \left(b + \frac{c}{2}\right)\alpha_2 .\end{aligned}$$

Now this tells us that $\beta$ is in $\langle\alpha_1, \alpha_2\rangle$. That is, because $\alpha_3$ is linearly dependent on $\alpha_1$ and $\alpha_2$, *every* vector in $\langle\alpha_1, \alpha_2, \alpha_3\rangle$ must be in $\langle\alpha_1, \alpha_2\rangle$. In symbols

$$\langle\alpha_1, \alpha_2, \alpha_3\rangle \subseteq \langle\alpha_1, \alpha_2\rangle.$$

Conversely, every vector in $\langle\alpha_1, \alpha_2\rangle$ has the form

$$a\alpha_1 + b\alpha_2 = a\alpha_1 + b\alpha_2 + 0\alpha_3$$

and so must be in $\langle\alpha_1, \alpha_2, \alpha_3\rangle$; that is

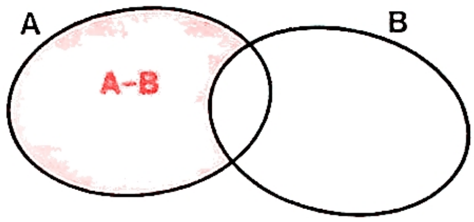
$$\langle\alpha_1, \alpha_2\rangle \subseteq (\alpha_1, \alpha_2, \alpha_3\rangle.$$

Putting these two inclusions together, we get

$$\langle\alpha_1, \alpha_2, \alpha_3\rangle = \langle\alpha_1, \alpha_2\rangle.$$

What we have illustrated here is the statement of ***Theorem 2.5*** on page N13, which you should note. The notation $A - \{\alpha_k\}$, used in that theorem,

means the set $A$ with the element $\alpha_k$ removed. In general if $A$ and $B$ are two sets, $A - B$ is the set $A$ with those elements removed which are also in $B$.




**Exercises**

1. Describe in geometrical terms the set spanned by each of the following subsets of vectors in $R^2$:

   (i) $\{(1, 0)\}$ (answer: the $x$-axis)
   (ii) $\{(0, 1)\}$
   (iii) $\{(1, 0), (0, 1)\}$
   (iv) $\{(1, 0), (2, 0)\}$
   (v) $\{(0, 0)\}$
   (vi) $\emptyset$ (the empty set)

2. In the space $P_3$ consisting of all real polynomials in $x$ of degree 2 or less (together with the zero polynomial), which of the following statements are true?

   (i) $x^2 - x - 1 \in \langle x^2, x + 1\rangle$
   (ii) $x^2 - x + 1 \in \langle x^2, x + 1\rangle$
   (iii) $\langle x^2, x + 1\rangle = P_3$
   (iv) $\langle x^2, x, 1\rangle = P_3$
   (v) $\langle x^2, x^2 - x, x, 1\rangle = P_3$
   (vi) $\langle x^2 - x, x, 1\rangle = P_3$

**Solutions**

1. (i) The $x$-axis
   (ii) The $y$-axis
   (iii) The whole plane
   (iv) The $x$-axis
   (v) The origin
   (vi) The origin (see page N12, line $-7$)

2. (i) True, since $x^2 - x - 1 = 1x^2 + (-1)(x + 1)$.
   (ii) False, since every polynomial in $\langle x^2, x + 1\rangle$ is of the form $\alpha x^2 + \beta(x + 1)$, and $x^2 - x + 1$ is not of this form.
   (iii) False, since the result of (ii) shows that $x^2 - x + 1$ is in $P_3$ but not in $\langle x^2, x + 1\rangle$.
   (iv) True, since the general element of $P_3$ can be written as $\alpha x^2 + \beta x + \gamma$, which is a linear combination of $x^2$, $x$, and 1.
   (v) True, since $\{x^2, x, 1\}$ spans $P_3$ and adding more vectors to the set cannot decrease $\langle x^2, x, 1\rangle$. Also, adding the vector $x^2 - x$ to $\{x^2, x, 1\}$ cannot increase $\langle x^2, x, 1\rangle$ since $x^2 - x \in \langle x^2, x, 1\rangle$.
   (vi) True, the general element of $P_3$ can be written

$$\alpha x^2 + \beta x + \gamma = \alpha(x^2 - x) + (\alpha + \beta)x + \gamma 1$$

and is therefore a linear combination of $x^2 - x$, $x$, and 1. This is an example of *Theorem 2.5* with $A = \{x^2, x^2 - x, x, 1\}$ and $\alpha_k = x^2$, since $x^2 = (x^2 - x) + x$.

## 1.2.4 The Replacement Theorem

In this section we discuss ***Theorem 2.7*** on page N13. This theorem is probably the most important in this unit because from it we can arrive at the concept of dimension for a vector space, as we shall see in the next sections. Because of its importance, and the importance of the technique used to prove it we will work through the proof in detail.

*READ the statement of **Theorem 2.7** on page* N13.

Before working through the proof let us try to see what the theorem is saying. One way of doing this is to think of an example where this theorem could apply. We want a vector space $V$ and a finite set of vectors which spans $V$. The example we have considered so far is $R^2$, and we saw in the previous section that $R^2$ is spanned by $\{\alpha_1, \alpha_2\}$ where $\alpha_1 = (1, 1)$ and $\alpha_2 = (3, -1)$ i.e., $R^2 = \langle\alpha_1, \alpha_2\rangle$.

So we have a vector space $R^2$ and a finite set $\{\alpha_1, \alpha_2\}$ which spans $R^2$; what then does the theorem tell us about the vector space $R^2$? The theorem asserts that any linearly independent subset of $R^2$ can contain at most 2 elements. Putting this another way, if we pick any subset $W$ of $R^2$ which contains *more* than 2 elements, then $W$ is a linearly dependent set. For example, in the preceding sub-section we saw that $\{(1, 1), (3, -1), (1, -1)\}$ was linearly dependent in $R^2$. ***Theorem 2.7*** tells us this immediately. Surely this is a most surprising statement about a vector space to deduce from such an innocent assumption about a set of vectors spanning the space!

Let us consider another example, the vector space $P_3$ of polynomials of degree at most 2. We can pick out the polynomials $p_1 = x^2$, $p_2 = x$, $p_3 = 1$; then it is easy to see that $\langle p_1, p_2, p_3\rangle = P_3$ because if $p = ax^2 + bx + c$ is any vector in $P_3$, then we have $p = ap_1 = bp_2 + cp_3$ and so $p$ is in $\langle p_1, p_2, p_3\rangle$. Thus, $P_3$ is a vector space spanned by a finite set of vectors containing 3 elements and so the theorem tells us that any linearly independent subset of $P_3$ contains at most 3 vectors: any 4 vectors in $P_3$ will be linearly dependent.

What about a proof of this theorem? The method of proof that we are going to give here is essentially that given by N. The method of proof in fact explains why it is called the *Replacement Theorem*.

We first look at the information we are given. We have a vector space $V$ and a finite set of vectors $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ which spans $V$. We are also given a subset of $V$, $B$ say, which is linearly independent, that is there is no non-trivial linear combination of the vectors in $B$ which gives us the zero vector. ($B$, of course, cannot contain the zero vector.) We suppose $B = \{\beta_1, \beta_2, \ldots\}$ where the dots indicate that $B$ may have any number of elements. This is all the information we are given

$$A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\},\ n \text{ finite};\ A \text{ spans } V$$
$$B = \{\beta_1, \beta_2, \ldots\};\ B \text{ is linearly independent, } B \subset V.$$

To prove the theorem we go through a "replacement process"; we replace each $\alpha$ by a $\beta$ in turn in such a way that at each stage we still have a set

which spans $V$. This replacement process needs justification, but schematically it looks like this:

$$\boxed{A = \{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n\}} \qquad A \text{ spans } V$$

$$\downarrow \quad \text{replace } \alpha_1 \text{ by } \beta_1$$

$$\boxed{A_1 = \{\beta_1, \alpha_2, \alpha_3, \ldots, \alpha_n\}} \qquad A_1 \text{ spans } V$$

$$\downarrow \quad \text{replace } \alpha_2 \text{ by } \beta_2$$

$$\boxed{A_2 = \{\beta_1, \beta_2, \alpha_3, \ldots, \alpha_n\}} \qquad A_2 \text{ spans } V$$

$$\downarrow \quad \text{replace } \alpha_3 \text{ by } \beta_3$$

Suppose in fact that we do have some way of doing this replacement process in such a way that each of the sets $A_1, A_2, \ldots$ still spans $V$; then two possibilities can occur.

(i) Either the set $B$ contains $n$ or fewer vectors, in which case our replacement process will end with no $\beta$s left over, or

(ii) we reach $A_n = \{\beta_1, \beta_2, \ldots, \beta_{n-1}, \beta_n\}$; each $\alpha$ has been replaced and there are still some $\beta$s left over.

In case (i) everything is fine. $B$ contains at most $n$ vectors in the first place for this case to occur, which is just what we wanted to prove.

In case (ii) there is at least one $\beta$ left over, say $\beta_{n+1}$. Then $A_n = \{\beta_1, \beta_2, \ldots, \beta_n\}$ spans $V$ and $\beta_{n+1} \in V$, so $\beta_{n+1}$ is a non-trivial linear combination of $\{\beta_1, \beta_2, \ldots, \beta_n\}$. But this contradicts $B$ being linearly independent! This means that if $B$ is linearly independent, case (ii) cannot occur and this replacement process *must* end in case (i). That is, $B$ contains at most $n$ vectors.

So the whole proof rests on this replacement process, on being able to go from $A$ to $A_1$ to $A_2$ and so on, up to $A_n$, and at each stage having a set which spans $V$. Note how important it is that we *do* have sets which span $V$: the fact that $A_n$ spans $V$ enables us to eliminate case (ii).

Let us look more closely and see how we can do this replacement process. We look at the first step. We have $A$ spanning $V$ and we want to replace an $\alpha$ by a $\beta$ in such a way that the new set, $A_1$, spans $V$.

We first add $\beta_1$ to $A$ giving $\{\beta_1\} \cup A = \{\beta_1, \alpha_1, \alpha_2, \ldots, \alpha_n\}$. Since $A$ spans $V$, $\beta_1$ is linearly dependent on $A$. So by ***Theorem 2.5***

$$\langle \{\beta_1\} \cup A \rangle = \langle \{\beta_1\} \cup A - \{\beta_1\} \rangle = \langle A \rangle.$$

Hence, since $A$ spans $V$

$$\langle \{\beta_1\} \cup A \rangle = V.$$

We now want to remove an $\alpha$ from $\{\beta_1\} \cup A$ to obtain a set which still spans $V$. To apply ***Theorem 2.5*** we must make sure that the $\alpha$ we remove is linearly dependent on the other elements of $\{\beta_1\} \cup A$. Since $A$ spans $V$ and $\beta_1 \neq 0$, we have

$$\beta_1 = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$$

where not all the $a_i$ are zero. Renumbering the vectors in $A$ if necessary so that $a_1 \neq 0$, we have

$$\frac{1}{a_1}\beta_1 = \alpha_1 + \frac{a_2}{a_1}\alpha_2 + \cdots + \frac{a_n}{a_1}\alpha_n.$$

Rearranging

$$\alpha_1 = \frac{1}{a_1}\beta_1 - \left(\frac{a_2}{a_1}\alpha_2 + \cdots + \frac{a_n}{a_1}\alpha_n\right),$$

i.e. $\alpha_1$ is linearly dependent on the other vectors in $\{\beta_1\} \cup A$. Thus by ***Theorem 2.5***

$$\begin{aligned}\langle\{\beta_1\} \cup A\rangle &= \langle\{\beta_1\} \cup A\rangle - \{\alpha_1\}\\ &= \langle\beta_1, \alpha_2, \alpha_3, \ldots, \alpha_n\rangle\\ &= \langle A_1\rangle.\end{aligned}$$

Thus we have shown first that

$$\langle\{\beta_1\} \cup A\rangle = V$$

and secondly that

$$\langle\{\beta_1\} \cup A\rangle = \langle A_1\rangle.$$

Putting these together, we have

$$\langle A_1\rangle = V$$

which is just what we wanted to show. We have replaced one $\alpha$ by a $\beta$ and obtained a set of vectors $A$, which spans $V$. This is at least the first stage in the chain of replacements successfully completed.

What about the next stage in the replacement process? We begin by adding $\beta_2$ to $A_1$ and, by the same argument as before ($\beta_2$ is linearly dependent on $A_1$), we have

$$\langle\{\beta_2\} \cup A_1\rangle = V$$

by ***Theorem 2.5.***

Now we want to show that one of the $\alpha$s can be removed from $\{\beta_2\} \cup A_1$ in such a way that the new set spans $V$.

Since $A_1$ spans $V$ and $\beta_2 \in V$

$$\beta_2 = b_1\beta_1 + b_2\alpha_2 + b_3\alpha_3 + \cdots + b_n\alpha_n$$

where not all the $b_i$ are zero.

Now if $b_2 = b_3 = \cdots = b_n = 0$, then $\beta_2$ is a linear combination of $\beta_1$, and $B = \{\beta_1, \beta_2, \beta_3, \ldots\}$ would not have been linearly independent—contrary to our original statement. Thus, one of $b_2, b_3, \ldots, b_n$ is non-zero, and again renumbering $\alpha_2, \alpha_3, \ldots, \alpha_n$ if necessary, we may take $b_2 \neq 0$.

Thus

$$\frac{1}{b_2}\beta_2 = \frac{b_1}{b_2}\beta_1 + \alpha_2 + \frac{b_3}{b_2}\alpha_3 + \cdots + \frac{b_n}{b_2}\alpha_n$$

so that

$$\alpha_2 = \frac{1}{b_2}\beta_2 - \left(\frac{b_1}{b_2}\beta_1 + \frac{b_3}{b_2}\alpha_3 + \cdots + \frac{b_n}{b_2}\alpha_n\right).$$

We can now go through the same argument as in the first replacement process, using ***Theorem 2.5*** again to give

$$\begin{aligned}\langle\{\beta_2\} \cup A_1\rangle &= \langle\{\beta_2\} \cup A_1 - \{\alpha_2\}\rangle\\ &= \langle\beta_1, \beta_2, \alpha_3, \alpha_4, \ldots, \alpha_n\rangle\\ &= \langle A_2\rangle.\end{aligned}$$

Since we have already shown that

$$\langle\{\beta_2\} \cup A_1\rangle = V$$

we have

$$\langle A_2 \rangle = V.$$

So again we have been successful; the vector we are now calling $\alpha_2$ has been replaced by $\beta_2$ and we still have a spanning set for $V$.

Well, we just go through this process again and again, at each stage possibly renumbering the $\alpha$s that are left, but at each stage getting a set which spans $V$. So we see that the replacement of $\alpha$s by $\beta$s on which the proof of the theorem rests can be justified. There are two important points about this theorem: one is the theorem itself which we shall use in Section 1.3; the other is this technique of replacement which we shall use later to derive some important consequences. If you have time, try to review the proof of this theorem by reading the one given in **N** (page N13).

**Exercise**

The vector $\alpha = (3, 2)$ is linearly independent in $R^2$. Use the replacement process on the spanning set $\{(1, 0), (0, 1)\}$ for $R^2$ to find another vector $\beta \in R^2$, which together with $\alpha$ spans $R^2$.

**Solution**

Following the replacement process we wish to take out one of $(1, 0)$ and $(0, 1)$ from the set

$$\{\alpha, (1, 0), (0, 1)\}$$

such that the set remaining spans $R^2$. We do this by using the fact that $\{(1, 0), (0, 1)\}$ spans $R^2$ and hence write $\alpha$ as a linear combination of $\{(1, 0), (0, 1)\}$

$$\alpha = (3, 2) = 3(1, 0) + 2(0, 1).$$

Solving for $(0, 1)$ gives

$$(0, 1) = \frac{\alpha}{2} - \frac{3}{2}(1, 0).$$

Thus

$$\langle \alpha, (1, 0), (0, 1) \rangle = \langle (1, 0), (0, 1) \rangle = R^2$$

and

$$\langle \alpha, (1, 0) \rangle = \langle \alpha, (1, 0), (0, 1) \rangle.$$

Hence

$$\langle \alpha, (1, 0) \rangle = R^2$$

i.e.

$$\beta = (1, 0).$$

Alternatively, we could have written

$$(1, 0) = \frac{\alpha}{3} - \frac{2}{3}(0, 1)$$

and arrived at

$$\beta = (0, 1).$$

In this exercise, one could obviously solve the problem very easily by many other means. But that is not the point; suppose a similar problem were posed in $P_4$, say. The method remains the same: we have an effective algorithm for its solution.

### 1.2.5 Summary of Section 1.2

In this section we defined the terms

| | | |
|---|---|---|
| linear combination | (page N11) | * * * |
| linear relation | (page N11) | * * * |
| linearly dependent | (page N11) | * * * |
| linearly independent | (page N11) | * * * |
| span | (page N12) | * * * |

We introduced the notation $\langle A \rangle$ for the set spanned by the set $A$.

We considered three theorems in particular:

1. (*2.1*, page N12)
   If $\alpha$ is linearly dependent on $\{\beta_i\}$ and each $\beta_i$ is linearly dependent on $\{\gamma_j\}$ then $\alpha$ is linearly dependent on $\{\gamma_j\}$. * *
2. (*2.5*, page N13)
   If $\alpha_k \in A$ is dependent on the other vectors in $A$ then * *
   $$\langle A \rangle = \langle A - \{\alpha_k\} \rangle.$$
3. (*2.7*, page N13)
   The Replacement Theorem. If a finite set $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ spans $V$ then every linearly independent subset of $V$ contains at most $n$ elements. * * *

We studied the replacement process used to prove the Replacement Theorem. * *

## 1.3 BASES

### 1.3.0 Introduction

In the present section we use the ideas of linear dependence and span, together with the replacement theorem which connects them, to arrive at a definition of the *dimension* of a vector space. This idea of dimension considerably simplifies some of the results about linear dependence and span which we obtained in Section 1.2. We know already that the space we live in has 3 dimensions, that a plane has 2 dimensions and a line 1, but these ideas are not very precise, and it is not clear at this stage whether the idea of dimension applies to all vector spaces or whether it is a particular property of geometrical spaces. By defining dimension in a way that depends directly on the axioms of a vector space, we shall show that the concept of number of dimensions applies to *all* vector spaces.

The argument proceeds from the assumption that the vector space contains a set of vectors which is both linearly independent and a spanning set. Such a set is called a *basis*, and the number of elements in it is called the *dimension* of the space. The main thing we have to prove, to justify calling this the dimension, is that all bases in a given vector space contain the same number of elements. Thus, the number of elements in any basis is a characteristic of the space itself and does not depend on the particular basis chosen. In fact this number, the dimension, may be said to determine completely the structure of the vector space; for we shall see that any two vector spaces with the same dimension over the same field have precisely the same structure—i.e. there is an isomorphism (structure-preserving mapping) connecting them. For example, the geometric vectors in a plane and the complex numbers are both real vector spaces of dimension 2, and so there must be an isomorphism between them.

### 1.3.1 Definition of Basis

*READ Section* 3 *on page* N15 *as far as but not including* ***Theorem 3.1.***

**Notes**

(i) *line 15, page* N15 Example (1) is on page N8. The notation

$$\{\alpha_i = x^i \mid i = 0, 1, \ldots\}\ .$$

means the set whose elements are

$$\alpha_0 = 1, \qquad \alpha_1 = x, \qquad \alpha_2 = x^2, \ldots;$$

i.e., the basis is the set $\{1, x, x^2, \ldots\}$.
(ii) *line 17, page* N15 The definition of $P_n$ is in Example (2) on page N8.
(iii) *line 21, page* N15 This is just a compact way of saying that $R^2$ has the basis $\{\alpha_1 = (1, 0),\ \alpha_2 = (0, 1)\}$; that $R^3$ has the basis $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ and that, in general, $R^n$ has a basis where the $i$th basis vector $\alpha_i$ has zero entries everywhere except in the $i$th place where the entry is 1. In the case of $R^2$ the proof that $(1, 0)$ and $(0, 1)$ span the space is that any vector $(a, b)$ in $R^2$ can be written in the form $a(1, 0) + b(0, 1)$. The proof of independence is that

$$c(1, 0) + d(0, 1) = (0, 0)$$

implies

$$(c, d) = (0, 0)$$

i.e.

$$c = d = 0.$$

The proof for $R^3$, etc. is similar. The Kronecker delta occurs frequently in mathematics: it is worth getting used to it.

There are three important points in this reading passage. The first is that a basis *spans* the vector space; the second is that a basis is a *linearly*

*independent* set of vectors, and thirdly that every element in a vector space has a *unique* representation as a linear combination of basis vectors. We shall return to this third point in sub-section 1.3.3.

*Example*

As opposed to exercises which are meant for you to work before looking at the solutions, examples are meant to illustrate the methods and ideas in the current section. But, if you are doing well with the unit, you can turn an example into an exercise by just reading the statement of the problem and then trying it for yourself before reading the solution.

Let $p_1 = x^2 + 1$, $p_2 = x + 1$, $p_3 = x^2 + x + 1$ be three vectors in $P_3$, the vector space of real polynomials of degree at most 2. Show that $A = \{p_1, p_2, p_3\}$ is a basis for $P_3$.

*Solution*

In order to do this, we have to show that (i) $A$ is linearly independent and (ii) that $\langle A \rangle = P_3$.

(i) The standard method of proving a set of vectors linearly independent is to show that the only linear relation between the given vectors is the trivial relation with all the scalars zero.

Let the linear relation be

$$ap_1 + bp_2 + cp_3 = 0 \text{ (the zero polynomial)}$$

where $a$, $b$, $c$ are real numbers. Then

$$\begin{aligned} &a(x^2 + 1) + b(x + 1) + c(x^2 + x + 1) \\ &\quad = (a + c)x^2 + (b + c)x + (a + b + c)1 \\ &\quad = 0 \end{aligned}$$

This implies that $a + c = 0$,

$$b + c = 0$$

and $a + b + c = 0$.

The only solution to these three equations is

$$a = b = c = 0;$$

so that $A$ is linearly independent.

(ii) By the definition of $P_3$ we have

$$P_3 = \langle 1, x, x^2 \rangle,$$

and using the method of the Replacement Theorem we can replace $\{1, x, x^2\}$ by the linearly independent set $\{p_1, p_2, p_3\}$ to obtain $P_3 = \langle p_1, p_2, p_3 \rangle = \langle A \rangle$.

### 1.3.2 Dimension

This sub-section deals with *Theorems 3.1* to *3.6* on pages N15–17.

Let us suppose that we have a vector space $V$ and two bases, $A$ and $B$, for it; then can we find any connection between $A$ and $B$? In the last example we saw that $P_3$ has two bases, $\{1, x, x^2\}$ and $\{x^2 + 1, x + 1, x^2 + x + 1\}$, and we notice that both bases have the same number of elements. Again $\{(1, 0), (0, 1)\}$ and $\{(1, 1), (3, -1)\}$ are bases for $R^2$, as we saw in sub-section 1.2.2, and both bases have the same number of elements. This is no accident, as the next theorem, *Theorem 3.1*, shows. The proof of *Theorem 3.1* follows immediately from *Theorem 2.7* and it will help to clarify *Theorem 3.1* if, after reading the statement of the theorem, you try to write down your own proof of *Theorem 3.1*.

*READ from Theorem 3.1 on page* N15 *as far as (but not including) Theorem 3.2 on page* N16.

**Notes**

(i) *line* $-7$, *page* N15 "... the dimension ... is well defined." What is meant here is that we have defined the word *dimension* in terms of the number of elements in a basis for the vector space, and if we did not have a theorem like *3.1*, then we could not talk about *the* dimension of a vector space, but only about a dimension relative to a given basis.
(ii) *lines* $-6$ *to* $-3$, *page* N15 These statements are just definitions of convenience, as you will see later.
(iii) *line* $-3$, *page* N15 Many of the vector spaces whose elements are functions are infinite-dimensional and, as we shall see in later units, they are of very great interest indeed. It is worth while, in looking at any theorem about vector spaces, to see whether the theorem applies to arbitrary vector spaces. If it is restricted to the finite-dimensional case one should consider the reason why. An interesting case of this was *Theorem 2.7*, where the spanning set A has to be finite, otherwise the replacement process would go on for ever.

*Theorem 3.1* tells us a lot about bases for finite-dimensional vector spaces but it does not help us *find* a basis for any particular space. However, there are ways of finding bases, at least for finite-dimensional spaces, and many of them rely on the replacement technique. We list some of them below and indicate briefly how one might use the Replacement Theorem to prove that they work. Remember that a basis has two essential properties:

(i) it spans the vector space;
(ii) its vectors are linearly independent.

The first two theorems below show that if we know the vector space is $n$-dimensional and we have a set of $n$ elements, then we can assert that these $n$ elements are a basis if we have *either* (i) *or* (ii): we don't need both.

*Theorem* A set $B$ of $n$ linearly independent vectors in an $n$-dimensional vector space $V$ forms a basis for $V$.

To prove this theorem we want to show that $B$ spans $V$. The proof follows from observing that $V$ is $n$-dimensional and so must have a basis $A$ containing $n$ vectors. Now a basis spans $V$, so $\langle A \rangle = V$; if we use the replacement technique to replace the elements of $A$ by those in $B$, the process will end with all the vectors in $A$ having been replaced by those in $B$. That is, $B$ also spans $V$ and hence is a basis for $V$.

The converse of our theorem is also true; any basis for $V$ is a set of $n$ linearly independent vectors. Our theorem and its converse together form *Theorem 3.3* on page N16.

In a similar way we can show:

***Theorem*** A set $A$ of $n$ vectors which spans an $n$-dimensional vector space $V$ is also a basis for $V$.

To prove this theorem we want to show that $A$ is a linearly independent set. We use a proof by contradiction: suppose the vectors in $A$ are linearly dependent; then we can discard at least one to get a set $A_1$ of $n-1$ or fewer linearly independent vectors which spans $V$ (this is the result stated as ***Theorem 2.5*** in N). But this would mean that $V$ is of dimension $n-1$ or less, which contradicts the definition of $V$. Hence our supposition is false: the set $A$ is *not* linearly dependent and the theorem is proved.

The converse of this theorem is also true: any basis for $V$ is a set of $n$ vectors which spans $V$. Our theorem and its converse together form ***Theorem 3.4*** on page N16.

Finally, here is another theorem which is an immediate consequence of the replacement process and the fact that a finite-dimensional vector space has a finite spanning set.

***Theorem 3.6*** (page N17) In a finite-dimensional vector space $V$ any linearly independent set of vectors $B$ can be extended to a basis.

The idea behind this proof of this is easy. Take any basis $A$ for $V$ and use the replacement process of ***Theorem 2.7*** to insert all the elements of $B$ in $A$, getting rid of an equal number of elements of $A$ from the set as we proceed.

The following example illustrates the use of this theorem in finding a basis.

*Example*

The set $\{\beta = (1, 3)\}$ is linearly independent in $R^2$; so the theorem tells us it can be extended to give a basis. Although, in this simple case, it is very easy to complete the basis (any vector which is not a scalar multiple of $(1, 3)$ will do), we apply the theory explicitly, to illustrate the method for less obvious cases.

If $\alpha_1 = (1, 0)$ and $\alpha_2 = (0, 1)$,

$$\{\alpha_1, \alpha_2\} \text{ spans } R^2;$$

therefore $\{\beta, \alpha_1, \alpha_2\}$ spans $R^2$.

Having added $\beta$, we now wish to eliminate either $\alpha_1$ or $\alpha_2$ to complete the replacement process.

Since $\beta \in R^2$ and $\langle \alpha_1, \alpha_2 \rangle = R^2$,

$\beta$ is a linear combination of $\alpha_1$ and $\alpha_2$:

$$\beta = \alpha_1 + 3\alpha_2$$

Hence

$$\alpha_1 = \beta - 3\alpha_2 .$$

So we can replace $\alpha_1$ by $\beta$ to obtain the set $\{\beta, \alpha_2\}$, which also spans $R^2$. Since $R^2$ is of dimension 2, any set of 2 elements which spans $R^2$ must be a basis. So $\{\beta, \alpha_2\}$ is a basis containing $\beta$.

**Exercise**

1. Construct a basis for the real vector space $P_3$ which includes the vector $p = x^2 - 3$.
2. Page N19, Exercise 2. (The vector space is $R^3$, which is 3-dimensional.)

**Solution**

1. We use the replacement process. $\{1, x, x^2\}$ is a basis for $P_3$ and so spans $P_3$.

   The set $\{p, 1, x, x^2\}$ is linearly dependent; in fact

   $$p = -3 \cdot 1 + 1 \cdot x^2$$

   so

   $$1 = -\tfrac{1}{3}p + \tfrac{1}{3}x^2.$$

   Thus the set $\{p, x, x^2\}$ also spans $P_3$, and since $P_3$ has dimension 3 this set is a basis for $P_3$ (***Theorem 3.4***). The replacement process does not, of course, necessarily give a unique result. For instance, in this case, $\{p, 1, x\}$ is also a basis; the replacement process allows us to replace either 1 or $x^2$, but not $x$.

2. To show that

   $$(1, 0, 0) \in \langle(1, 1, 0), (1, 0, 1), (0, 1, 1)\rangle$$

   we obtain

   $$(1, 0, 0) = \tfrac{1}{2}(1, 1, 0) + \tfrac{1}{2}(1, 0, 1) - \tfrac{1}{2}(0, 1, 1).$$

   Similarly

   $$(0, 1, 0) = \tfrac{1}{2}(1, 1, 0) - \tfrac{1}{2}(1, 0, 1) + \tfrac{1}{2}(0, 1, 1)$$
   $$(0, 0, 1) = -\tfrac{1}{2}(1, 1, 0) + \tfrac{1}{2}(1, 0, 1) + \tfrac{1}{2}(0, 1, 1)$$

   Now any vector in $R^3$ is of the form

   $$(a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1)$$

   which by the above three relations is a linear combination of $(1, 1, 0)$, $(1, 0, 1)$ and $(0, 1, 1)$. Thus, we have shown that $\{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ spans $R^3$. Since $R^3$ is 3-dimensional, it follows by ***Theorem 3.4*** that this set is a basis.

### 1.3.3 Coordinates

This sub-section deals with the remainder of Section 3 on pages N17–19.

In this sub-section, we see how we can use a basis for a finite-dimensional vector space $V$ to give a unique numerical representation of every vector in $V$, by means of coordinates, which are analogous to the Cartesian coordinates used in analytic geometry. This will make it possible to give a complete description of a finite-dimensional vector space over a field $F$. To see how coordinates are introduced, let us consider the real vector space $P_3$ of all polynomials of degree at most 2. We have seen that the set $\{1, x, x^2\}$ forms a basis for $P_3$ and that any polynomial $p = a + bx + cx^2$ is expressed as a linear combination of these basis vectors:

$$p = a1 + bx + cx^2$$

We noted at the beginning of Section 3 that this representation for $p$ as a linear combination of $1, x, x^2$ is unique. That is, given the polynomial $p$, the coefficients $a, b, c$ are uniquely determined. We can use this uniqueness of representation to introduce coordinates into $P_3$. Given the basis $\{1, x, x^2\}$, any polynomial $p = a + bx + cx^2$ is determined uniquely by the coefficients $a, b, c$; so in order to specify $p$ we only need to specify the ordered triple $(a, b, c)$. The elements of this triple are called the *coordinates* of $p$ with respect to the basis $\{1, x, x^2\}$. The last part of this last sentence emphasizes that the coordinates on their own are not enough. The same coordinates with a different basis give a different vector; for example, if the basis were $\{1, x, x^2 - x\}$ instead of $\{1, x, x^2\}$, then the polynomial specified by the ordered triple of numbers $(1, 2, 3)$ would be

$$1 + 2x + 3(x^2 - x) = 1 - x + 3x^2$$

instead of

$$1 + 2x + 3x^2.$$

*READ from page* N17 *line −11, to the end of the section on page* N19 (*omitting the exercises*).

**Notes**

*line 13, page* N18 In the Foundation Course we paid considerable attention to the idea of a *morphism*, by which we mean, roughly, a mapping that preserves some mathematical structure. An *isomorphism* is a morphism with an inverse which is also a morphism, and the domain and codomain of an isomorphism are said to be *isomorphic*. Here we are concerned with the structure described by the vector space axioms.

**Exercise**

Find the coordinates of

$$p = x^2 - 3x + 2$$

with respect to the basis $\{x - 1, x + 1, x^2\}$.

**Solution**

If $(a, b, c)$ are the coordinates of $p$ with respect to the given basis, then

$$\begin{aligned} p &= cx^2 + b(x + 1) + a(x - 1) \\ &= cx^2 + (b + a)x + (b - a)1 \end{aligned}$$

But

$$p = x^2 - 3x + 2$$

and $\{x^2, x, 1\}$ is also a basis. Therefore the linear combination of $x^2$, $x$ and 1 which gives $p$ is unique. So

$$c = 1, \qquad b + a = -3, \qquad b - a = 2$$

i.e. the coordinates of $p$ are $(-\frac{5}{2}, -\frac{1}{2}, 1)$ with respect to

$$\{x - 1, x + 1, x^2\}.$$

### 1.3.4 Summary of Section 1.3

In this section we defined the terms

| | | |
|---|---|---|
| basis | (page N15) | * * * |
| dimension | (page N15) | * * * |
| coordinates | (page N17) | * * * |

We discussed the following theorems

1. (*3.1*, page N15)
   If a vector space has one basis with a finite number of elements, then all other bases are finite and have the same number of elements. * * *
2. (*3.3*, page N16)
   A set of $n$ vectors in an $n$-dimensional vector space $V$ is a basis if and only if it is linearly independent. * * *
3. (*3.4*, page N16)
   A set of $n$ vectors in an $n$-dimensional vector space $V$ is a basis if and only if it spans $V$. * * *
4. (*3.6*, page N17)
   In a finite-dimensional vector space any linearly independent set of vectors can be extended to a basis. * * *

## 1.4 SUBSPACES

### 1.4.0 Introduction

A vector space $V$, with its field $F$, is an example of a mathematical structure; that is, a set with certain algebraic operations defined on the set. Fields and groups are two other examples of mathematical structures which you have already met.

Whenever we want to find out more about a structure, there are two things we can do. One is to look at the structure as a whole, by comparing it with other similar structures: the mathematical instrument we use to do this is the morphism. We have already used this instrument to a considerable degree in the Foundation Course, and we shall be using it again in later units of this course. Also the idea of coordinates as we have just seen, uses the morphism idea in the form of an isomorphism. The other way of finding out more about a structure, which is itself often interlinked with the first, is to take the structure to pieces as it were, by looking for simpler versions of the same structure within itself: that is, we look for subsets of the original set which have the same structure. In the case of a vector space, this means that we look for subsets of the vector space which are themselves vector spaces. Geometrically, this may be visualized as studying three-dimensional space by looking at *some* of the planes and straight lines in it; since every vector space must have a zero element, we are restricted when studying such subspaces to planes and lines *through the origin*.

We shall not take this study very far at this stage: if you are interested and have the time, you may like to read more of Section 4 of N than we indicate in the following. We shall merely define a subspace and discuss criteria for a subset of a vector space to be a subspace.

### 1.4.1 Definition of a Subspace

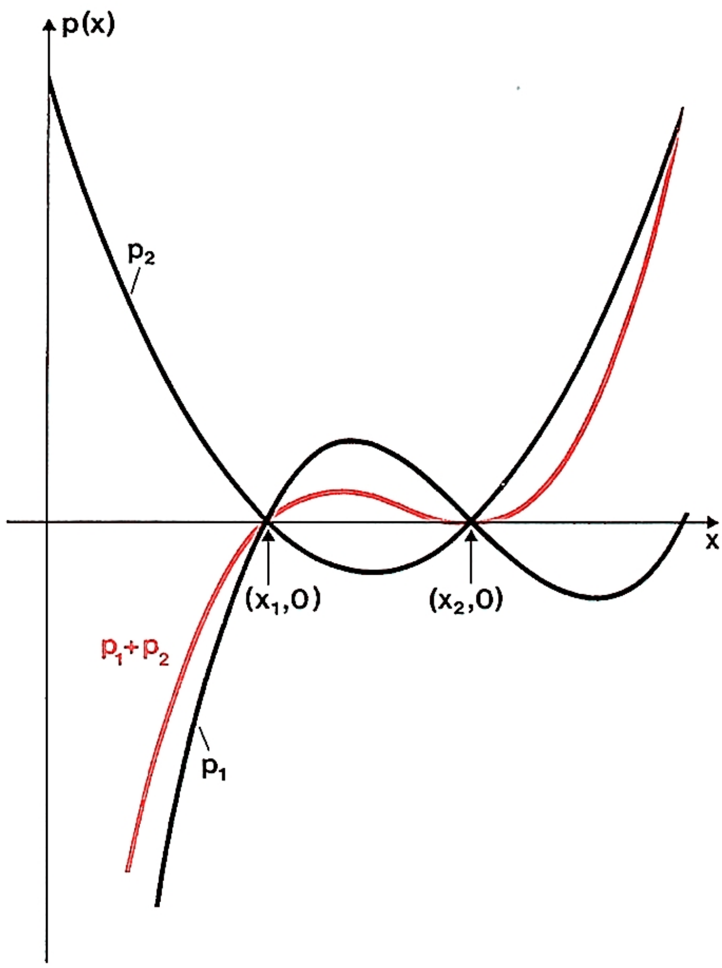
*READ Section 4 starting on page* N20 *as far as (but not including)* ***Theorem 4.1*** *on page* N21.

**Notes**

(i) *line* $-6$, *page* N20 The notation $p(x_1)$, where $p$ is a polynomial and $x_1$ is a number, means the image of $x_1$ under the corresponding polynomial function. For example, if $p = 2x + 3$ then $p(x_1) = 2x_1 + 3$, and $p(10) = 23$.
(ii) *line* $-3$, *page* N20 If $p_1$, $p_2$ are polynomials such that $p_1(x_1) = p_2(x_1) = 0$, then

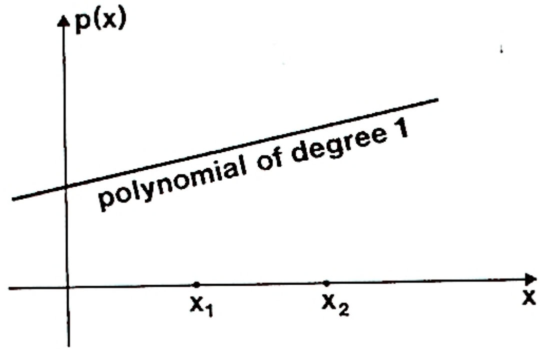
$$\begin{aligned}(p_1 + p_2)(x_1) &= p_1(x_1) + p_2(x_1)\\ &= 0 + 0\\ &= 0\end{aligned}$$

so that $p_1 + p_2$ also vanishes at $x_1$; and the same is true of $x_2$, $x_3$, etc. Similarly if $a$ is a real number, then $(ap_1)(x) = a(p_1(x)) = a0 = 0$.



(iii) *line* $-2$, *page* N20 If $m > n$, then we have polynomials of degree $n < m$ with $m$ zeros. (N.B. A polynomial $p$ has a zero at $x$ if $p(x) = 0$.) There is a general theorem in algebra which states that the only such polynomial is the zero polynomial. (In fact, $m = n$ would give the same result, since $P_n$ consists of all polynomials of degree less than or equal to $n - 1$.) Thus the answer to N's question is that we still get a subspace, but it is trivial, i.e. it contains the zero polynominal only.

One can visualize the theorem for $P_2$. If we suppose that a polynomial $p$ of degree 1 has two zeros, $x_1$ and $x_2$, then $p(x_1) = p(x_2) = 0$. But geometrically $p$ is represented by a straight line, which must pass through $(x_1, 0)$ and $(x_2, 0)$. It is therefore the $x$-axis, i.e. $p$ is the zero polynomial.




There are two important points to note in this passage. The first is that all you have to do to check that a subset $W$ of a vector space $V$ is a subspace is to show

(a) $W$ is non-empty; that is, there is at least one vector in $W$ (which may just be the zero vector);
(b) if $\alpha_1, \alpha_2 \in W$ and $a, b \in F$, then $a\alpha_1 + b\alpha_2 \in W$; i.e. $W$ is closed under the taking of linear combinations.

We shall use these two properties over and over again in proving theorems and statements about subspaces. The other point is that we have an important property of a spanning set, one which we did not discuss in Section 1.2, namely that, if $A$ is a subset, $\langle A \rangle$ is a *subspace*.

**Exercises**

1. Page N24, Exercise 1.
2. Page N25, Exercise 11.
3. Answer N's "Why?" on line 7, page N21.

**Solutions**

1. The answers to this exercise are given in N. We add notes on the first and last part.

   (a) This is a subspace. Call the subset $X$; we have to verify two things. First $X$ is non-empty; the zero polynomial is in $X$ since it has the value 0 everywhere, in particular at 1. That is, $X$ is non-empty. Secondly, $X$ is closed under the taking of linear combinations. We consider the polynomial

   $$q = ap_1 + bp_2$$

   where $a$ and $b$ are real numbers, and $p_1$ and $p_2$ are in $X$. To see if $q$ is in $X$, we calculate $q(1)$.

   $$\begin{aligned} q(1) &= (ap_1 + bp_2)(1) \\ &= (ap_1)(1) + (bp_2)(1) \\ &= a(p_1(1)) + b(p_2(1)) \end{aligned}$$

$$= a \cdot 0 + b \cdot 0 \quad \text{since we assumed that } p_1, p_2 \in X$$
$$= 0 + 0$$
$$= 0.$$

That is, $q(1) = 0$ and $q$ *is* in $X$.

Hence we have shown that $X$ is a subspace of $P$.

(e) The subset here is not a subspace, since, for example,

$$(x^2 + 2x - 1) + (-x^2 + x + 2) = 3x + 1$$

and the latter is not of even degree.

2. There are two parts to the answer. We have to show that

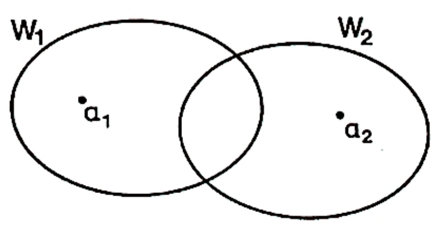
(i) if $W_1$ and $W_2$ are subspaces such that one is a subspace of the other, then $W_1 \cup W_2$ is a subspace

(ii) if $W_1$ and $W_2$ are subspaces such that neither one is a subspace of the other, then $W_1 \cup W_2$ is not a subspace.

N's answer concerns (ii) only.

(i) is easily established. If, for example, $W_1 \subset W_2$, then $W_1 \cup W_2 = W_2$, and hence $W_1 \cup W_2$ is a subspace.

(ii) The following diagram may help you to understand N's answer, which we paraphrase below.




Consider

$$\alpha_1 \in W_1 - W_2$$

and

$$\alpha_2 \in W_2 - W_1$$

Then, since $\alpha_1 \in W_1$ and $\alpha_2 \notin W_1$,

$$\alpha_1 + \alpha_2 \notin W_1$$

for otherwise, $\alpha_2$ would be in $W_1$.

Similarly

$$\alpha_1 + \alpha_2 \notin W_2.$$

Since

$$\alpha_1 + \alpha_2 \notin W_1$$

and

$$\alpha_1 + \alpha_2 \notin W_2,$$
$$\alpha_1 + \alpha_2 \notin W_1 \cup W_2.$$

Thus $W_1 \cup W_2$ is not a subspace, since it is not closed.

As an example, consider the subspaces of $P$ given in (a) and (b) of Exercise 1.

$$\alpha_1 = x - 1 \text{ belongs to the first subspace}$$
$$\alpha_2 = x^2 + 2x \text{ belongs to the second subspace}$$

but $\alpha_1 + \alpha_2 = x^2 + 3x - 1$ belongs to neither subspace and so does not belong to their union.

3. The subset of $R^n$ of all $n$-tuples of rational numbers, $Q^n$, is not a subspace over $R$ since it is not closed under scalar multiplication. For example

$$q = (\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \ldots, \tfrac{1}{2}) \in Q^n$$

but

$$\pi q = \left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}, \ldots, \frac{\pi}{2}\right) \notin Q^n.$$

## 1.4.2 Summary of Section 1.4

In this section we defined the term

subspace (page N20) * * *

We introduced a criterion for deciding if a subset of a vector space is a subspace: * * *

a subset $W$ of a vector space $V$ is a subspace, if

(a) $W$ is non-empty

and if

(b) $\alpha_1, \alpha_2 \in W$ and $a, b \in F$, then $a\alpha_1 + b\alpha_2 \in W$; i.e. $W$ is closed under the taking of linear combinations.

## 1.5 SUMMARY OF THE UNIT

### Definitions

The terms defined in this unit and references to their definitions are given below.

A *field* is any mathematical structure satisfying the axioms $F1$ to $F11$ on page N6–7. * * *

A *vector space* (over some given field $F$) is any mathematical structure satisfying the axioms $A1$ to $A5$ and $B1$ to $B5$ on pages N7–8. In particular, the set is closed under the operations of addition, and multiplication by any scalar (i.e. element of $F$). * * *

You are not expected to memorize these lists of axioms, but you should be able to use them to check whether a given structure is, or is not, a vector space.

Other terms defined are

| | | |
|---|---|---|
| linear combination | (page N11) | * * * |
| linear relation | (page N11) | * * * |
| linearly dependent | (page N11) | * * * |
| linearly independent | (page N11) | * * * |
| span | (page N12) | * * * |
| basis | (page N15) | * * * |
| dimension | (page N15) | * * * |
| coordinates | (page N17) | * * * |
| subspace | (page N20) | * * * |

### Theorems

We list the important theorems discussed in this unit. Only 2- and 3-star theorems which are essential to this and later units have been included in this list. References to the statement of the theorems in N are also given.

1. (*2.1*, page N12)
   If $\alpha$ is linearly dependent on $\{\beta_i\}$ and each $\beta_i$ is linearly dependent on $\{\gamma_j\}$ then $\alpha$ is linearly dependent on $\{\gamma_j\}$. * *
2. (*2.5*, page N13)
   If $\alpha_k \in A$ is dependent on the other vectors in $A$ then * *
   $$\langle A \rangle = \langle A - \{\alpha_k\} \rangle$$
3. (*2.7*, page N13)
   The Replacement Theorem. If a finite set $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ spans $V$ then every linearly independent subset of $V$ contains at most $n$ elements. * * *
4. (*3.1*, page N15)
   If a vector space has one basis with a finite number of elements, then all other bases are finite and have the same number of elements. * * *
5. (*3.3*, page N16)
   A set of $n$ vectors in an $n$-dimensional vector space $V$ is a basis if and only if it is linearly independent. * * *
6. (*3.4*, page N16)
   A set of $n$ vectors in an $n$-dimensional vector space $V$ is a basis if and only if it spans $V$. * * *
7. (*3.6*, page N17)
   In a finite-dimensional vector space any linearly independent set of vectors can be extended to a basis. * * *

### Techniques

1. The replacement process, which tells us how to include a set of linearly independent vectors in a spanning set. * *

2. A criterion to decide if a subset is a subspace. * * *

A subset $W$ of a vector space $V$ is a subspace, if

(a) $W$ is non-empty

and if

(b) $\alpha_1, \alpha_2 \in W$ and $a, b \in F$, then $a\alpha_1 + b\alpha_2 \in W$; i.e. $W$ is closed under the taking of linear combinations.

**Notation**

$\langle A \rangle$ denotes the set spanned by the set of vectors $A$.

**Examples**

We introduced three important examples of vector spaces; all were first introduced in sub-section 1.1.1.

$R^n$ The space of $n$-tuples of real numbers * * *

$$(a_1, a_2, \ldots, a_n),\ a_i \in R.$$

$C[a, b]$ The space of all real functions continuous on the interval $[a, b]$. * * *

$P$ The real vector space of polynomials with real coefficients. * * *

## 1.6 SELF-ASSESSMENT

### Self-assessment Test

This Self-assessment Test is designed to help you test quickly your understanding of the unit. It can also be used, together with the summary of the unit, for revision. The answers to these questions will be found on the next non-facing page. We suggest you complete the whole test before looking at the answers.

In the following questions choose the most appropriate option from those given to complete the sentence.

**1.** The set of vectors $\{(1, 1), (2, 2), (e, e), (\pi, \pi)\}$ is linearly ____B____ in $R^2$.

A independent
B dependent

**2.** The set of vectors $\{(1, 1), (3, -1), (1, -1)\}$ spans a/an ____B____ dimensional vector space.

A 1
B 2
C 3
D infinite

**3.** The set of vectors $W = \{(0, a, b): a, b \in R\}$ forms a ____C____ $R^3$.

A basis for
B spanning set for
C subspace of

**4.** The polynomials $1 + x$, $1 - x$ form a ________ $P_2$.

A subspace of
B basis for
C subset of

Classify each of the following statements as either true or false.

**5.** In any $n$-dimensional vector space, any subset of $n$ vectors spans $V$. F
**6.** Any linearly independent subset of $V$ can be extended to a basis for $V$. T
**7.** Any set which spans $V$ is a basis for $V$. F

The following questions require some calculation.

**8.** The set $\{1 + x^2, 1 - x, 1 + x + x^2, x - x^2\}$ is a spanning set for $P_3$. Use the Replacement Theorem to include $\{1 - x^2, 1 + x\}$ in a spanning set for $P_3$.
**9.** Specify the set spanned by $\{(1, 1, 0), (1, 0, 1)\}$ in $R^3$.

## Solutions to Self-assessment Test

| | | |
|---|---|---|
| **1.** | B, dependent. | Each vector is a linear combination of $(1, 1)$. |
| **2.** | B, 2. | The vectors form a subset of $R^2$ so that C and D are obviously wrong answers. On the other hand $(1, 1)$ and $(1, -1)$ are linearly independent so that $\langle(1, 1), (1, -1)\rangle = R^2$; so A is also wrong. |
| **3.** | C, subspace. | $W$ is an infinite subset of $R^3$, a space of dimension 3; so A is wrong, any basis for $R^3$ must have just 3 elements. $W$ does not span $R^3$ because $(1, 0, 0) \notin \langle W\rangle$. |
| **4.** | B, basis. | B is the most appropriate answer, while C is also true. $\{1 + x, 1 - x\}$ is not a subspace; it does not, for example, contain the zero vector. |
| **5.** | False. | The $n$ vectors may be linearly dependent. |
| **6.** | True. | See sub-section 1.3.2, ***Theorem 3.6***. |
| **7.** | False. | See sub-section 1.3.2, ***Theorem 3.4***. Also question 2 above. |

**8.** To make the use of the Replacement Theorem more obvious, let

$$\alpha_1 = 1 + x^2, \qquad \alpha_2 = 1 - x, \qquad \alpha_3 = 1 + x + x^2,$$
$$\alpha_4 = x - x^2,$$
$$\beta_1 = 1 - x^2, \qquad \beta_2 = 1 + x.$$

We first note that $\beta_1, \beta_2$ are linearly independent so we can use the Replacement Theorem.

If we add $\beta_1$ to $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, which $\alpha$ can be omitted? We write $\beta_1$ as a linear combination of the $\alpha$s.

$$\begin{aligned}\beta_1 &= 1 - x^2\\ &= 1 - x + x - x^2\\ &= \alpha_2 + \alpha_4.\end{aligned}$$

So

$$\alpha_2 = -\alpha_4 + \beta_1$$

and we replace $\alpha_2$ by $\beta_1$.

$P_3$ is spanned by $\{\beta_1, \alpha_1, \alpha_3, \alpha_4\}$.

Now write $\beta_2$ in terms of this spanning set.

$$\begin{aligned}\beta_2 &= 1 + x\\ &= 1 + x + x^2 - x^2\\ &= 1 + x^2 + x - x^2\\ &= \alpha_1 + \alpha_4.\end{aligned}$$

So

$$\alpha_1 = \beta_2 - \alpha_4$$

and we replace $\alpha_1$ by $\beta_2$.

This gives $\{\beta_1, \beta_2, \alpha_3, \alpha_4\}$ as a spanning set for $P_3$.

**9.** The set spanned by $\{(1, 1, 0), (1, 0, 1)\}$ is the set of all vectors of the form

$$a(1, 1, 0) + b(1, 0, 1) \qquad a, b \in R$$

i.e.

$$\{(a + b, a, b): \qquad a, b \in R\}.$$

If we think of the geometric representation of this set, then $(x, y, z)$ belongs to the set if

$$x = y + z$$

and this is a plane through the origin.

LINEAR MATHEMATICS

1 Vector Spaces
2 Linear Transformations
3 Hermite Normal Form
4 Differential Equations I
5 Determinants and Eigenvalues
6 NO TEXT
7 Introduction to Numerical Mathematics: Recurrence Relations
8 Numerical Solution of Simultaneous Algebraic Equations
9 Differential Equations II: Homogeneous Equations
10 Jordan Normal Form
11 Differential Equations III: Nonhomogeneous Equations
12 Linear Functionals and Duality
13 Systems of Differential Equations
14 Bilinear and Quadratic Forms
15 Affine Geometry and Convex Cones
16 Euclidean Spaces I: Inner Products
17 NO TEXT
18 Linear Programming
19 Least-squares Approximation
20 Euclidean Spaces II: Convergence and Bases
21 Numerical Solution of Differential Equations
22 Fourier Series
23 The Wave Equation
24 Orthogonal and Symmetric Transformations
25 Boundary-value Problems
26 NO TEXT
27 Chebyshev Approximation
28 Theory of Games
29 Laplace Transforms
30 Numerical Solution of Eigenvalue Problems
31 Fourier Transforms
32 The Heat Conduction Equation
33 Existence and Uniqueness Theorem for Differential Equations
34 NO TEXT

The Open University

*Mathematics: A Second Level Course*

*Linear Mathematics Unit 2*

# LINEAR TRANSFORMATIONS

*Prepared by the Course Team*

The Open University Press

The Open University Press Walton Hall Milton Keynes MK7 6AA

First published 1971. Reprinted 1975, 1976


Designed by the Media Development Group of the Open University.

Printed in Great Britain by
Martin Cadbury

SBN 335 01091 1

This text forms part of the correspondence element of an Open University Second Level Course. The complete list of units in the course is given at the end of this text.

For general availability of supporting material referred to in this text, please write to the Director of Marketing, The Open University, P.O. Box 81, Walton Hall, Milton Keynes, MK7 6AT.

Further information on Open University courses may be obtained from the Admissions Office, The Open University, P.O. Box 48, Walton Hall, Milton Keynes, MK7 6AB.

1.3

## Contents

## Set Books

D. L. Kreider, R. G. Kuller, D. R. Ostberg and F. W. Perkins, *An Introduction to Linear Analysis* (Addison-Wesley, 1966).

E. D. Nering, *Linear Algebra and Matrix Theory* (John Wiley, 1970).

It is essential to have these books; the course is based on them and will not make sense without them.

## Conventions

Before working through this correspondence text make sure you have read *A Guide to the Linear Mathematics Course*. Of the typographical conventions given in the Guide the following are the most important.

The set books are referred to as:

> **K** for *An Introduction to Linear Analysis*
> N for *Linear Algebra and Matrix Theory*

All starred items in the summaries are examinable.

References to the Open University Mathematics Foundation Course Units (The Open University Press, 1971) take the form *Unit M100 3, Operations and Morphisms*.

## 2.0 INTRODUCTION

You have met linear transformations previously if you took the Foundation Course, M100. They are functions that map one vector space to another; in other words, they are vector space morphisms (or homomorphisms, as they are called in N). One example of a linear transformation, which we saw in *Unit 1, Vector Spaces,* is the mapping from any real vector space $V$ of dimension $n$ to the space $R^n$, in which each vector in $V$ is mapped to its coordinates with respect to some chosen basis in $V$.

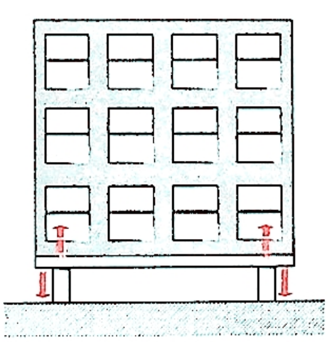
In geometry, the linear transformations* include rotations about the origin, reflections in a line or plane through the origin, and dilations (magnifications), or in general, any mapping that maps parallel straight lines to parallel straight lines and the origin to the origin. We shall meet many other examples of linear transformations during the course.

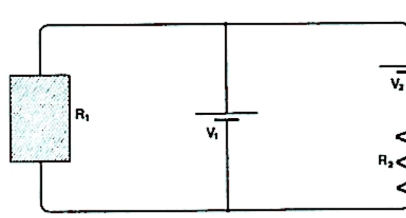
We study transformations partly for their own interest, partly to help us understand the structure of vector spaces (by mapping them to other spaces, such as $R^n$, or a geometrical space, whose structure is easier to understand), and partly because they crop up so often when we try to solve problems using mathematics. You will have met many such problems already; for example, a system of linear simultaneous algebraic equations such as

$$\begin{aligned} 2x - 3y &= 3 \\ x + y &= 5 \end{aligned}$$

can be considered as arising from a linear transformation of $R^2$ to itself. We discussed this in some detail in *Units M100 23* and *26*, *Linear Algebra II* and *III*. As further examples consider the following. (You are not expected to understand them fully; their purpose is to show you that the subject does have applications.)

(1) The reinforced concrete beam in the building below is subjected to the forces indicated. What would be the deformations if the forces were changed?






(2) Two batteries of voltages $V_1$ and $V_2$, the second with a large resistance $R_2$, drive a circuit element of arbitrary nature, except that it has resistance $R_1$. The problem is to determine the currents in the circuit.

* Some authors of geometry texts use a different definition from the one we shall use in this course. So, if you venture into geometry, check definitions first.

(3) A periodic driving force is applied to a mechanical system that can vibrate. How does it move? (We discussed some aspects of this problem in *Unit M100 31, Differential Equations II.*)

(4) Suppose that a survey revealed that 3% of country-dwellers move into town each year and 1% of town-dwellers move into the country each year. At present, 50% of the population live in towns. What is the future population distribution going to be?

How do we build mathematical models of these situations? At this stage, we have to decide what are the basic features of the system we are considering. In fact, all the above examples have been chosen to illustrate the same basic features: one aspect or *state* of the system is related to another aspect or *state*; i.e. there is a transformation between the characteristics of one state and those of the other. Furthermore the relation can be shown to be linear. The following table shows in the second and third columns the related states for each example.

| | | | |
|---|---|---|---|
| (1) | Building | Stresses | Deformations |
| (2) | Circuit | Voltages | Currents |
| (3) | Oscillating system | Driving force | Motion |
| (4) | Society | Population Distribution now | Population Distribution in the future. |

If you appreciate that each of these problems involves a relationship between the characteristics of one state and those of another and that this relationship is independent of the particular way we choose to represent these states, then you will see that it makes sense in the mathematics to consider linear transformations in the abstract without regard to any specific numerical representation.

In Section 2.1, we find numbers that characterize a linear transformation in the same way that the dimension characterizes a vector space; these numbers are the dimensions of the two vector spaces concerned (the domain and image set), and a number called the nullity of the transformation.

In Section 2.2, we look at the representation of a linear transformation by scalars (the *matrix* representation), analogous to the representation of a vector by coordinates.

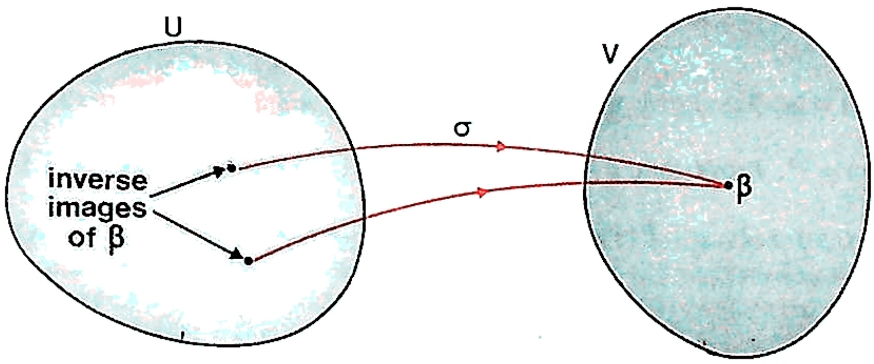
Consider the inverse problem to that of Example 3 above. If we know how the mechanical system moves, can we determine the periodic driving force? Section 2.3 looks at the special matrix properties involved in such inverse problems.

The material in this unit is essential to all the subsequent units. The difficulty and the amount of time that it will take you will depend on your recall and grasp of the linear algebra units of the Foundation Course, as well as the units on mappings and morphisms. If you find this unit and *Unit 1* taking longer than you would like, it is probably because they contain a large amount of revision material.

## 2.1 LINEAR TRANSFORMATIONS

### 2.1.1 The Definition of a Linear Transformation

Chapter II, Section 1 of Nering, which starts on page N27, defines a linear transformation, and gives names to various special types of linear transformation. The only definition that you have not met already in the Foundation Course is that of inverse image: if $\sigma$ is a linear transformation from a vector space $U$ to a vector space $V$, and $\beta = \sigma(\alpha)$ is the image of a vector $\alpha \in U$, then $\alpha$ is called an inverse image of $\beta$ under $\sigma$.

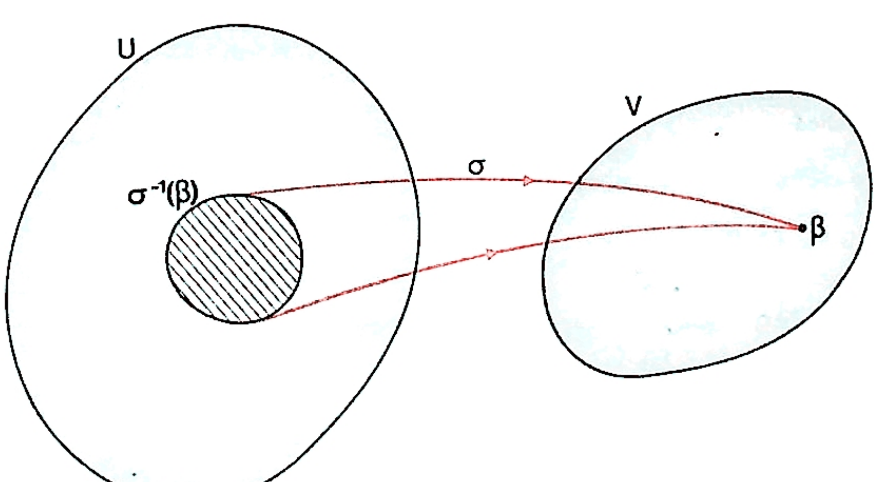



Of course, there may well be more than one inverse image of $\beta$ under $\sigma$. The collection of all the inverse images of $\beta$, which is a set, is called the *complete inverse image* of $\beta$. We know it from the Foundation Course as the image of $\beta$ under the *reverse* mapping of $\sigma$.

*READ Chapter II from the beginning on page* N26 *to the bottom of page* N27 (*excluding last two lines*).

### Notes

(i) *line 4, page* N26 In Foundation Course language we could phrase this "A linear transformation is a function whose domain is a vector space $U$ and whose codomain is the same or another vector space $V$, which is a morphism with respect to the taking of linear combinations."
(ii) *line −6, page* N26 "We identify $U$ and $V$" means "we treat $U$ and $V$ as identical", not "we establish the identity of $U$ and $V$".
(iii) *lines 1–7, page* N27 Hermite normal form is the subject of *Unit 3*.
(iv) *line 15, page* N27 "Single-valued mapping" is another way of saying "function".
(v) *line 20, page* N27 If $\beta$ is the image of $\alpha$, then $\alpha$ is an inverse image of $\beta$, and the set of all inverse images of $\beta$ is the complete inverse image, $\sigma^{-1}(\beta)$.



In the Foundation Course we described inverse images in terms of the *reverse mapping*, whose images are the inverse images for the original mapping.
In the Foundation Course you also met the *kernel* of the transformation $\sigma$, which is the complete inverse image of the zero vector in $V$. For example, if $U$ is a suitable set of functions and $\sigma$ is the differentiation function (or operator), then the function sin is an inverse image of the function cos, and the complete inverse image of cos is the set of all functions of the form

$$\sin + \text{constant function}$$

(vi) *line* $-6$, *page* N27 In the Foundation Course we used the word *morphism* in the way "homomorphism" is used here. In particular, this means that an isomorphism is a special case of a homomorphism.

**Exercises**

1. Exercise 1 on page N35.
2. Define $\sigma\colon R \longrightarrow R^2$ by $\sigma(x) = (x, 2x)$. Prove that $\sigma$ is a linear transformation of $R$ into $R^2$.
3. Define $\sigma\colon R \longrightarrow R$ by $\sigma(x) = x^2$. Prove that $\sigma$ is not a linear transformation.
4. Define $\sigma\colon R^2 \longrightarrow R^2$ by $\sigma((x, y)) = (2x, y^2)$. Prove that $\sigma$ is not a linear transformation.

**Solutions**

1. To show that the given $\sigma$ is a linear transformation we must show that it satisfies the definition on page N27. Since $R^2$ is a real vector space, the field called $F$ in this definition is $R$. $\sigma$, in this case, is clearly a function; so we are left to check that it satisfies Equation (1.1) on page N27. To calculate the left-hand side of the equation, we must specify $\alpha$ and $\beta$ as vectors in $R^2$.

   Let

   $$\alpha = (y_1, y_2),\ \beta = (z_1, z_2)$$

   so that

   $$a\alpha + b\beta = (ay_1 + bz_1, ay_2 + bz_2)$$

   (by rule (1.2) on page N9).

   Then the left-hand side of the equation in the definition is, by the definition of $\sigma$,

   $$\sigma(a\alpha + b\beta) = (ay_2 + bz_2, ay_1 + bz_1)$$

   The right-hand side is

   $$\begin{aligned} a\sigma(\alpha) + b\sigma(\beta) &= a(y_2, y_1) + b(z_2, z_1) \\ &= (ay_2 + bz_2, ay_1 + bz_1) \end{aligned}$$

   and since this equals the left-hand side for all $a, b, y_1, y_2, z_1, z_2$, the definition is satisfied.

2. Equation (1.1) on page N27 can be established as follows:

   $$\begin{aligned} \sigma(ax + by) &= (ax + by, 2(ax + by)) \\ &= (ax + by, 2ax + 2by) \\ &= (ax, 2ax) + (by, 2by) \\ &= a(x, 2x) + b(y, 2y) \\ &= a\sigma(x) + b\sigma(y). \end{aligned}$$

3. That Equation (1.1) does not hold for all $a$ and $b$ can be seen in the following. If $b = 0$, and $a$ is not 1 or 0, then

$$\begin{aligned}\sigma(ax+by) &= \sigma(ax)\\ &= a^2x^2\\ &= a^2\sigma(x)\\ &\neq a\sigma(x).\end{aligned}$$

4. If $b = 0$,

$$\begin{aligned}\sigma(a(x, y) + b(w, z)) &= \sigma(a(x, y))\\ &= \sigma((ax, ay))\\ &= (2ax, a^2y^2)\\ &= a(2x, ay^2).\end{aligned}$$

and $$\begin{aligned}a\sigma((x, y)) + b\sigma((w, z)) &= a\sigma((x, y))\\ &= a(2x, y^2).\end{aligned}$$

Thus, unless $a = 0$ or 1, Equation (1.1) does not hold.

## 2.1.2 Isomorphisms

One reason for studying morphisms is to compare different mathematical structures. In general, the structure in the image set of a morphism may be simpler than the one it came from in the domain: as an extreme example, there is a linear transformation that maps every element of the domain to the vector space consisting of the zero element only; thus it completely obliterates the original structure. We might call this a "forgetful" mapping! The other extreme is a mapping with a perfect memory. A morphism that does transfer the domain structure in full detail to the codomain is called an *isomorphism*, and can be characterized by the fact that it has an inverse, which is also a morphism, by which we can return from the codomain to the domain and recover the original structure in full detail. The next reading passage begins by considering the conditions under which a morphism has an inverse function (it must be one–one and onto) and then shows (***Theorem 1.1***) that this inverse function is an isomorphism. You should know the statement of this theorem and be able to follow the proof, but you are not expected to be able to reproduce the proof.

*READ from the last two lines on page* N27 *to the paragraph ending with the words "or a one-element set" near the foot of page* N28.

### Notes

(i) *line* $-1$, ***page*** N27 Ignore the term "monomorphism": the term *one-to-one* (or *one–one*) is enough for this course.
(ii) *line 3*, ***page*** N28 In the Foundation Course we used the term *image set* of $\sigma$, rather than "image of $\sigma$" for the set of all images under $\sigma$.
(iii) *line 4*, ***page*** N28 In words, *onto* means that each element of $V$ has at least one inverse image. Ignore the term "epimorphism": the term *onto* is sufficient for this course.
(iv) *lines 19–21*, ***page*** N28 For *an epimorphism*, read *onto* and for *a monomorphism*, read *one-to-one*. By notes (i) and (iii) above, if $\sigma$ is both one-to-one and onto, then each element of $V$ has no more and no less than one inverse image.
(v) *line 23*, ***page*** N28 The mapping $\sigma^{-1}$ is a function.
(vi) *The first line of* ***Theorem 1.1***, ***page*** N28 In this theorem, $\sigma$ stands for the isomorphism itself. Notice the strategy of proof: it is the same as that of Solution 1 on page 8. That is, we prove that $\sigma^{-1}$ satisfies Equation (1.1) on page N27.

*Example*

Show that $\sigma\colon R \longrightarrow R$, $\sigma(x) = 2x$ is an isomorphism. Find $\sigma^{-1}$.

*Solution*

*Linearity* $$\begin{aligned}\sigma(ax+by) &= 2(ax+by) = 2ax+2by\\ &= a\sigma(x) + b\sigma(y).\end{aligned}$$

*One–one* Suppose $x$ and $x'$ are two different real numbers.

Then $\sigma(x) = 2x$
and $\sigma(x') = 2x'$
i.e. $\sigma(x) \neq \sigma(x')$.

Thus $\sigma$ is one–one, by the definition at the foot of page N27.

*Onto* As $x$ runs through all of $R$, so does $2x$; i.e. given any $y \in R$, there is an element $\frac{1}{2}y$, such that $\sigma(\frac{1}{2}y) = y$. Thus $\sigma$ is an isomorphism since it is one–one and onto.

*Inverse* Clearly, $\sigma^{-1}: R \longrightarrow R$, $\sigma^{-1}(x) = \frac{1}{2}x$.

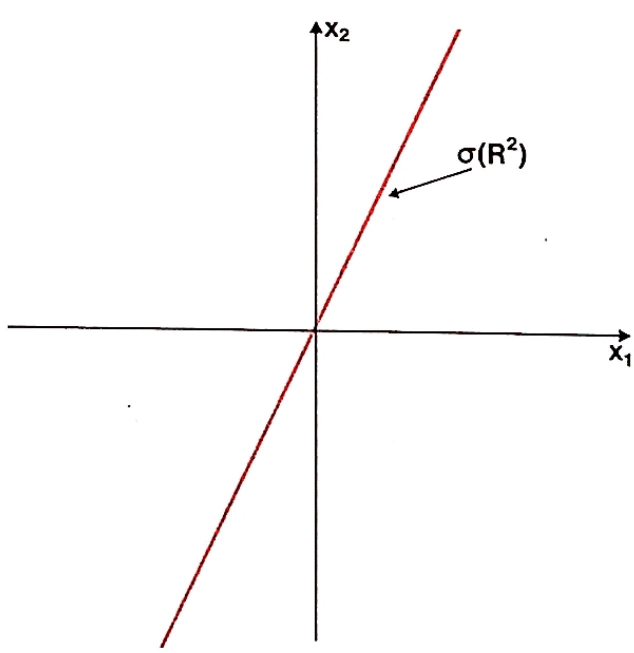
While it is possible to do the above mentally, there is a definite value in pedantically setting everything out formally until one is familiar with the process.

**Exercises**

1. Show that $\sigma: R^2 \longrightarrow R^2$, $\sigma((x_1, x_2)) = (x_1, 2x_1)$ is *not* an isomorphism.
2. Show that $\sigma((x_1, x_2)) = (2x_1, 2x_2)$ defines an isomorphism from $R^2$ to $R^2$. Find $\sigma^{-1}$ and verify that it satisfies the definition of an isomorphism.

**Solutions**

1. Clearly, $\sigma$ is a linear transformation, but that is not enough. In fact, let $x_2$, $x_2'$ be two distinct real numbers. Then
$\sigma((x_1, x_2)) = (x_1, 2x_1)$
and $\sigma((x_1, x_2')) = (x_1, 2x_1)$.
Thus, $\sigma$ is many–one. This is sufficient to show that $\sigma$ is not an isomorphism. But you may notice that it is also not "onto" $R^2$. The figure shows that the image set is the line whose equation is $x_2 = 2x_1$ and not the whole plane.




2. To satisfy the definition on page N28, we show that $\sigma$ is a homomorphism and is one–one and onto. It is a homomorphism (linear transformation) since

$$\begin{aligned}\sigma(a(x_1, x_2) + b(y_1, y_2)) &= \sigma((ax_1 + by_1, ax_2 + by_2)) \\ &= (2ax_1 + 2by_1, 2ax_2 + 2by_2)\end{aligned}$$

and $a\sigma((x_1, x_2)) + b\sigma((y_1, y_2)) = a(2x_1, 2x_2) + b(2y_1, 2y_2)$
$= (2ax_1 + 2by_1, 2ax_2 + 2by_2)$.

It is one–one and onto because the equation

$$\sigma((x_1, x_2)) = (y_1, y_2),$$

has exactly one solution, $(x_1, x_2)$, for each $(y_1, y_2)$. This solution is $(x_1, x_2) = (\frac{1}{2}y_1, \frac{1}{2}y_2)$ and so

$$\sigma^{-1}: (y_1, y_2) \longmapsto (\tfrac{1}{2}y_1, \tfrac{1}{2}y_2).$$

The verification that $\sigma^{-1}$ is an isomorphism is the same as for $\sigma$, but with $\frac{1}{2}$ replacing 2 everywhere.

### 2.1.3 Examples of Linear Transformations

*READ the last paragraph on page* N28 *and the first two paragraphs on page* N29 *finishing at the words "will denote the identity transformation on U".*

**Notes**

(i) *line* $-1$, *page* N28 "The indefinite integral" means the image under the one–many mapping:

$$f \longmapsto \text{(all primitive functions of } f)$$

which we denoted by $I$ in *Unit M100 13, Integration II.*
(ii) *line 1, page* N29 The exercise for this section will be to verify that $\sigma$ is onto, etc.
(iii) *line 13, page* N29 The second exercise in the previous section dealt with a scalar transformation $\sigma: (x_1, x_2) \longmapsto (2x_1, 2x_2)$.

**Exercises**

Verify the statements about $\sigma$ and $\tau$ in the first complete sentence on page N29.

**Solution**

The mapping $\sigma$ is the differentiation mapping for polynomials. It is *onto P*, since, given any polynomial

$$\alpha = \sum_{i=0}^{n} a_i x^i$$

there is always another polynomial

$$\beta = \sum_{i=0}^{n} \frac{a_i}{i+1} x^{i+1}$$

such that

$$\sigma(\beta) = \alpha.$$

In fact $\beta = \tau(\alpha)$. $\beta$ is only one of the polynomials whose image under $\sigma$ is $\alpha$. $\sigma$ is not one-to-one since two different polynomials can have the same derivative (if they differ by a constant).

The mapping $\tau$ maps a polynomial to a particular one of its indefinite integrals (primitives). It is one-to-one because two different polynomials (or functions) cannot have the same indefinite integral: if the image

$$\sum_{i=0}^{n} \frac{a_i}{i+1} x^{i+1}$$

is known, then all the coefficients, $a_0, a_1, a_2, a_3, \ldots, a_n$ in the original polynomial

$$\sum_{i=0}^{n} a_i x^i$$

are known.

$\tau$ is not onto, because only the polynomials having zero as their constant terms are images under $\tau$.

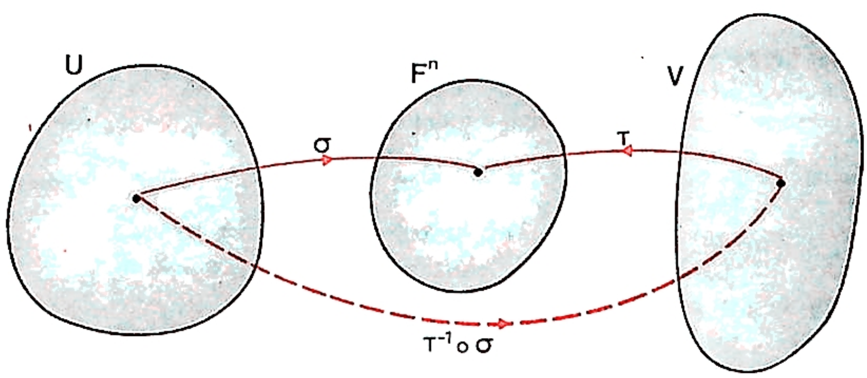
### 2.1.4 The Representation Isomorphisms

We have already mentioned, in *Unit 1*, *Vector Spaces*, the fact that any vector space of dimension $n$ over a field $F$ is isomorphic to $F^n$. It is an easy step from here to prove that any two vector spaces of the same finite dimension $n$ over the same field are isomorphic.

Thus if $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for one of the spaces, $U$ say, and $\sigma$ is the mapping from $U$ to $F^n$ defined by

$$\sigma: a_1\alpha_1 + \cdots + a_n\alpha_n \longmapsto (a_1, \ldots, a_n)$$

then $\sigma$ is an isomorphism (as shown at the top of page N18). If $\tau$ is the corresponding isomorphism taking the other space $V$ to $F^n$, then by performing first $\sigma$ and then $\tau^{-1}$ we obtain a mapping from $U$ to $V$ which is also an isomorphism.




The next reading passage discusses this important isomorphism briefly.

*READ* (*a*) *page* N29, *the one paragraph beginning* "*When a basis ...*".
(*b*) *page* N48 *the one paragraph beginning* "*Let U and V be vector spaces...*".

#### Notes

*Line 9 of the third paragraph on page* N29 It does not matter if you cannot see the point of the remarks about *natural* or *canonical* isomorphisms at this stage. Until we come to the unit on linear functionals, none of the isomorphisms we consider will be natural. The important thing is that the isomorphism from $V$ to $F^n$ considered here depends on the basis used in $V$: use a different basis and you get a different isomorphism.

## 2.1.5 The Combination of Linear Transformations

Since linear transformations are functions, they can be combined in various ways, as discussed in *Unit M100 1, Functions.*

*READ from the last paragraph on* N29 *as far as (but not including)* ***Theorem 1.2*** *on page* N31.

### Notes

(i) *line* $-8$, *page* N29 The addition of linear transformations gives a convenient notation for linear differential equations. For example, if $U$ is a suitable set of functions, and $D$ is the linear transformation of differentiation then we can define linear transformations such as

$$2D + 3 : f \longmapsto 2Df + 3f = 2f' + 3f \qquad (f \in U)$$

and write an equation such as

$$2f' + 3f = g,$$

where $g$ is some given function, in a form such as

$$(2D + 3)f = g.$$

The advantage of this notation is that one can specify the linear transformation in a very concise way, without using any arrows or "dummy variables". We shall use this notation in the units on differential equations.
(ii) *line 16, page* N30 The Foundation Course notation would be $\tau \circ \sigma$ rather than $\tau\sigma$, and instead of *iteration* or *multiplication*, we would have used *composition*.

The passage you have just read could be summarized as follows:

1. The set of all linear transformations with domain $U$ and codomain $V$, denoted in N by Hom $(U, V)$, is itself a vector space. Here, Hom $(U, V)$ stands for *homo*morphisms of $U$ into $V$.
2. Linear transformations can sometimes be "multiplied" together by the rule for composition of mappings; but this type of multiplication cannot be defined between *every* pair of mappings. Even when it is defined, it is not generally commutative. Note that these "multiplications" give Hom $(U, U)$ more structure than that of a vector space, since this operation of "multiplication" is in addition to those required to specify a vector space; i.e. vector addition and scalar multiplication.

### Exercises

1. Exercise 2 on page N35, taking the vector space in question to be $R^2$.
2. Prove that $\tau\sigma$, as defined on page N30, is a linear transformation.
3. Exercise 6 on page N36. (To translate this exercise into Foundation Course notation, it is easiest to take $y$ to be a function and to replace $\frac{dy}{dx}$ in the statement of the question by the derived function $y'$.) *Hint*: use the result of Exercise 2.

### Solutions

1. $(\sigma_1 + \sigma_2) : (x_1, x_2) \longmapsto (x_2, -x_1) + (x_1, -x_2)$
   i.e. $(\sigma_1 + \sigma_2) : (x_1, x_2) \longmapsto (x_1 + x_2, -x_1 - x_2)$.
   $\sigma_1\sigma_2 : (x_1, x_2) \longmapsto \sigma_1((x_1, -x_2))$
   i.e. $\sigma_1\sigma_2 : (x_1, x_2) \longmapsto (-x_2, -x_1)$.
   $\sigma_2\sigma_1 : (x_1, x_2) \longmapsto \sigma_2(x_2, -x_1)$
   i.e. $\sigma_2\sigma_1 : (x_1, x_2) \longmapsto (x_2, x_1)$.

   Note that $\sigma_2\sigma_1 \neq \sigma_1\sigma_2$; in fact $\sigma_1\sigma_2 = -\sigma_2\sigma_1$.

2. For any $a, b \in F$ and any $\alpha, \beta \in U$ we have

$$\begin{aligned}\tau\sigma(a\alpha + b\beta) &= \tau(a\sigma(\alpha) + b\sigma(\beta)) \text{ since } \sigma \text{ is linear}\\ &= a\tau(\sigma(\alpha)) + b\tau(\sigma(\beta)) \text{ since } \tau \text{ is linear}\\ &= a\tau\sigma(\alpha) + b\tau\sigma(\beta).\end{aligned}$$

   Thus $\tau\sigma$ is linear.

3. $D$ is linear since, for any two functions $y$ and $z$ in the domain of $D$,

$$D(ay + bz) = aD(y) + bD(z)$$

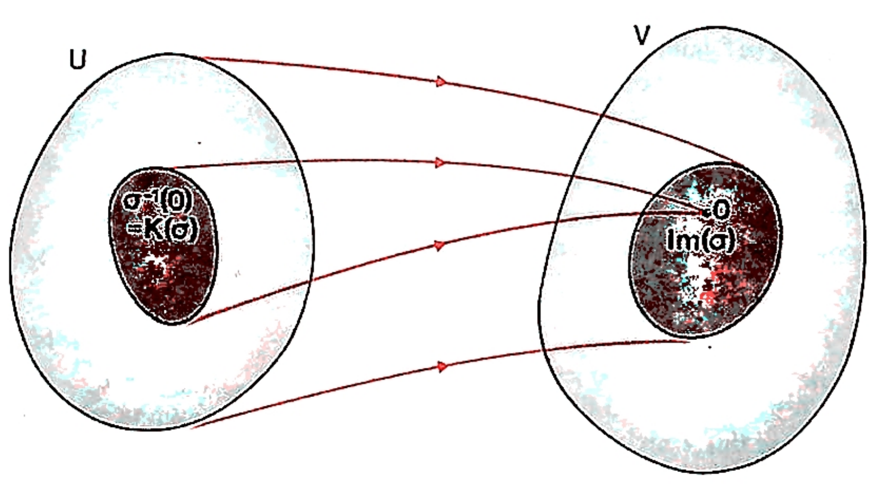
   by the rules of differentiation. By the result of Solution 2, it follows that $D^2 = DD$ is also linear, and hence that $D^3 = D^2D$ is also linear, and so on up to $D^n$. Further, if

$$p(D) = a_0 + a_1D + \cdots + a_nD^n,$$

   then this is the sum of the transformations $a_0, a_1D, \ldots, a_nD^n$, which are themselves linear, and hence $p(D)$ itself is linear by the result proved at the bottom of page N29.

## 2.1.6 Image Space and Kernel

Two subspaces are very naturally associated with every linear transformation. One of these subspaces is in the codomain, $V$, and is the image space denoted by $\mathrm{Im}(\sigma)$, consisting of all vectors that are images under $\sigma$. The other subspace is in the domain, $U$, and is our old friend from the Foundation Course, the kernel of $\sigma$. Denoted in N by $K(\sigma)$, it is the same as $\sigma^{-1}(0)$, the complete inverse image (as defined on page N27) of the zero element of $V$.




We first establish that $\mathrm{Im}(\sigma)$ and $K(\sigma)$ really are subspaces—that is, they are closed under the operation of taking any linear combination of two elements, and are not empty (see page N20). You have aready met the proofs that both of these sets are subspaces in the Foundation Course. To revise the proof that $\mathrm{Im}(\sigma)$ is a subspace

*READ **Theorem 1.2**, with its proof, on page* N31.

To revise the proof that $K(\sigma)$ is a subspace of $U$ we could use ***Theorem 1.5***. Instead we present the construction of this proof as a *programmed* exercise.

**Exercises**

1. Supply the missing symbols or groups of symbols in the following argument.

By definition, $K(\sigma)$ is the set

$$\{\alpha : \alpha \in U \text{ and } \boxed{\delta(\alpha)} = 0\}. \qquad \text{(i)}$$

To show that $K(\sigma)$ is a subspace we must show that, for all $a, b \in F$ and all $\alpha, \beta \in K(\sigma)$, we have

$$\boxed{a\alpha + b\beta} \in K(\sigma). \qquad \text{(ii)}$$

This is equivalent to showing that

$$\boxed{\delta(a\alpha + b\beta)} = 0. \qquad \text{(iii)}$$

But since $\sigma$ is a linear transformation, we have

$$\begin{aligned} \sigma(a\alpha + b\beta) &= \boxed{a\delta\alpha} + \\ &\quad \boxed{b\delta\beta} \qquad \text{(iv)} \\ &= 0 + 0 \qquad \text{since } \alpha, \beta \in K(\sigma) \\ &= 0 \end{aligned}$$

and the proof is complete.

Look back over your completed proof before you check it against the solution.

2. Exercise 3 on page N35. Do you notice anything about the dimensions of $\text{Im}(\sigma)$ and $K(\sigma)$?
3. Exercise 4 on page N35. ($\sigma$ is a mapping of $R^4$ to $R^2$.)

**Solutions**

1. (i) $\sigma(\alpha)$ (ii) $a\alpha + b\beta$ (iii) $\sigma(a\alpha + b\beta)$ (iv) $a\sigma(\alpha) + b\sigma(\beta)$.

2. $\text{Im}(\sigma)$ is the subset of $R^n$ defined by $\{(x_1, \ldots, x_k, 0, \ldots, 0) : x_1, \ldots, x_k \in R\}$. This subset is isomorphic to $R^k$ and so is $k$-dimensional. To describe $K(\sigma)$, we note that we need an $(x_1, \ldots, x_n)$ such that $\sigma(x_1, \ldots, x_n)$ is equal to $(x_1, \ldots, x_k, 0, \ldots, 0) = (0, \ldots, 0)$. But this can be true if and only if $x_1 = x_2 = \cdots = x_k = 0$.

   There are no further conditions restricting $x_{k+1}, \ldots, x_n$. Thus,

   $$K(\sigma) = \{(0, \ldots, 0, x_{k+1}, \ldots, x_n) : x_{k+1}, \ldots, x_n \in R\}.$$

   This subset is isomorphic to $R^{n-k}$ and is of dimension $n - k$. The notable fact about the dimensions of $\text{Im}(\sigma)$ and $K(\sigma)$ is that their sum is equal to $n$. This foreshadows ***Theorem 1.6***, which we also met in the Foundation Course.

3. We expect that you can show that $\sigma$ is a linear transformation without help, so we do not give a solution to this part. The determination of the kernel of $\sigma$ is equivalent to solving simultaneous linear equations, a technique you have met in the Foundation Course. This point is important to remember, as you will be required to find the kernel for various linear transformations and to relate this to solving a set of linear equations. Now by the definition of a kernel, we seek a $\xi \in R^4$ such that

   $$\sigma(\xi) = 0;$$

   i.e. if

   $$\xi = (x_1, x_2, x_3, x_4),$$

   then

   $$\sigma(\xi) = 0$$

is just

$$(3x_1 - 2x_2 - x_3 - 4x_4, x_1 + x_2 - 2x_3 - 3x_4) = (0, 0)$$

or

$$\begin{aligned} 3x_1 - 2x_2 - x_3 - 4x_4 &= 0 \\ x_1 + x_2 - 2x_3 - 3x_4 &= 0. \end{aligned}$$

This is a set of simultaneous linear equations as promised. In fact, there are 2 equations in 4 unknowns : thus, there are not enough equations to determine all of $x_1, \ldots, x_4$, but we can use them to determine 2 of the unknowns, say $x_1$ and $x_2$, in terms of the other 2. This gives

$$x_1 = x_3 + 2x_4, \; x_2 = x_3 + x_4.$$

This says that to be in the kernel of $\sigma$, the components of $\xi$ must be restricted as above, so that the kernel is

$$K(\sigma) = \{(x_3 + 2x_4, x_3 + x_4, x_3, x_4) : x_3, x_4 \in R\}.$$

### 2.1.7 Rank and Nullity

To characterize the subspaces $\mathrm{Im}(\sigma)$ and $K(\sigma)$ further, we investigate their dimensions. We define

$\rho(\sigma)$ = the *rank* of $\sigma = \dim \mathrm{Im}(\sigma)$

$\nu(\sigma)$ = the *nullity* of $\sigma = \dim K(\sigma)$.

As a minor point, note the following short paragraph from page N31. "For the rest of this book, unless specific comment is made, we assume that all vector spaces under consideration are finite dimensional. Let dim $U = n$ and dim $V = m$."

The most important result about rank and nullity is ***Theorem 1.6***: that their sum is equal to the dimension of the domain. We saw an example of this in an exercise in the preceding section. We shall call ***Theorem 1.6*** the *dimension theorem*, and it is the main result in Section 1 of this chapter of N. The dimension theorem was stated in *Unit M100 23, Linear Algebra II* and an optional proof was given in an appendix there. For the proof of the dimension theorem you can look back at *Unit M100 23*, read the proof in N, read our proof just below, or use the proof in the television programme. Do not be confused by the multiplicity of proofs; they all amount to the same thing, namely seeing what is left over in $U$ after you "hive off" $K(\sigma)$.

*Alternative Proof of Dimension Theorem :* $\rho(\sigma) + \nu(\sigma) = n$

$K(\sigma)$ is a subspace of $U$; that is, it is a vector space in its own right. This means that we can find a basis $\{\alpha_1, \ldots, \alpha_\nu\}$ for $K(\sigma)$, $\nu$ being short for $\nu(\sigma)$, the dimension of $K(\sigma)$. We know from ***Theorem 3.6*** on page N17 that we can extend this linearly independent set to give a basis

$$\{\alpha_1, \ldots, \alpha_\nu, \alpha_{\nu+1}, \ldots, \alpha_n\}$$

for $U$. We write

$$W = \langle \alpha_{\nu+1}, \ldots, \alpha_n \rangle.$$

We now have two subspaces, $W$ and $K(\sigma)$, of $U$. The most interesting thing to notice about these two subspaces is that

$$W \cap K(\sigma) = \{0\}.$$

We begin by proving this result. Let us assume that $\alpha \in W \cap K(\sigma)$. Then, (i) because $\alpha \in K(\sigma)$, we can write $\alpha$ as a linear combination of the basis vectors of $K(\sigma)$ :

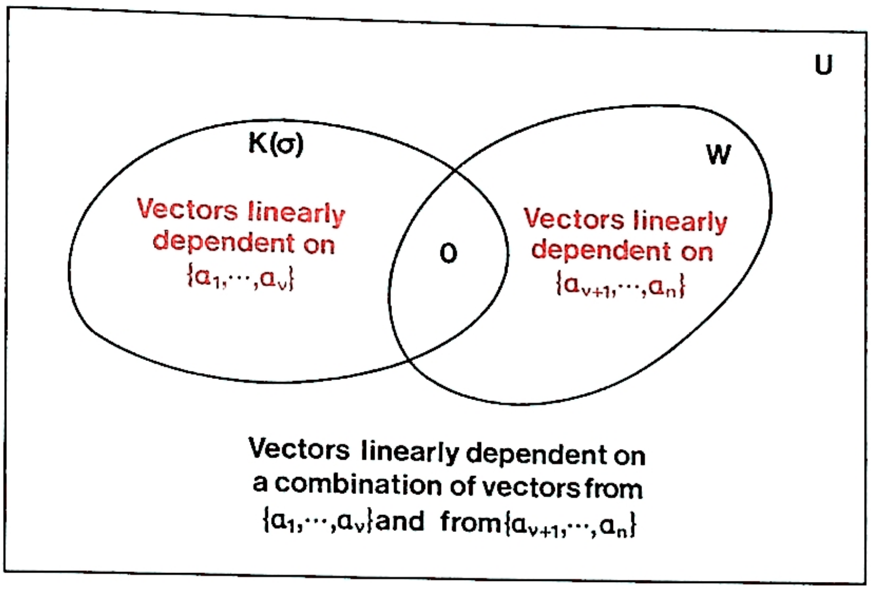
$$\alpha = c_1\alpha_1 + \cdots + c_\nu\alpha_\nu$$

and (ii) because $\alpha \in W$, we can write $\alpha$ as a linear combination of the basis vectors of $W$:

$$\alpha = c_{v+1}\alpha_{v+1} + \cdots + c_n\alpha_n.$$

But we know that $\alpha$ can be expressed uniquely in terms of a complete basis $\{\alpha_1, \ldots, \alpha_n\}$ of $U$, and we seem to have done it in two distinct ways. This contradiction can only be resolved if $\alpha = 0$.

So, we can now draw the following diagram.




This result, which is of some interest in its own right, is going to prove useful in the rest of the proof below.

$K(\sigma)$ has dimension $\nu(\sigma)$: the "complementary subspace" $W$ has dimension $n - \nu(\sigma)$, by its construction, and if we can prove that $W$ has the same dimension as $\text{Im}(\sigma)$, then we have proved that $\rho(\sigma) = n - \nu(\sigma)$, which is the dimension theorem. A way to prove that two spaces have the same dimension is to prove that they are isomorphic. To do this we shall show that the mapping $\sigma_1$ with domain $W$ (*not* $U$) and codomain $\text{Im}(\sigma)$ (*not* $V$) defined by

$$\sigma_1 : \alpha \longmapsto \sigma(\alpha) \qquad (\alpha \in W)$$

is an isomorphism. (We write $\sigma_1$ instead of $\sigma$ because we have changed domain and codomain; properly speaking we have a new mapping.) To show that $\sigma_1$ is an isomorphism we have to do three things:

(i) show that it is a linear transformation;
(ii) show that it is one–one;
(iii) show that it is onto (with codomain $\text{Im}(\sigma)$).

(i) is obvious; $\sigma_1$ is only $\sigma$ in disguise.
(ii) To show that $\sigma_1$ is one–one, we require to show that if $\alpha, \alpha' \in W$ and $\alpha \neq \alpha'$, then

$$\sigma_1(\alpha) \neq \sigma_1(\alpha').$$

Now since $\sigma_1$ is linear,

$$\sigma_1(\alpha) - \sigma_1(\alpha') = \sigma_1(\alpha - \alpha').$$

Since $W$ is a vector space,

$$\alpha - \alpha' \in W.$$

Also, since $W$ and $K(\sigma)$ have only zero in common and $\alpha \neq \alpha'$, $\alpha - \alpha' \notin K(\sigma)$;

i.e.

$$\sigma_1(\alpha - \alpha') \neq 0.$$

Thus

$$\sigma_1(\alpha) - \sigma_1(\alpha') \neq 0;$$

i.e.

$$\sigma_1(\alpha) \neq \sigma_1(\alpha').$$

So

$$\sigma_1 \text{ is one–one.}$$

(iii) Let $\beta \in \mathrm{Im}(\sigma)$. We want to show that there is an $\alpha \in W$ such that $\sigma_1(\alpha) = \beta$. We do know that there is an $\alpha \in U$ such that $\sigma(\alpha) = \beta$ because $\beta \in \mathrm{Im}(\sigma)$; let's see if we can prove that we can choose that $\alpha$ so that it belongs to $W$. There are two possibilities

(a) $\beta = 0$, in which case 0 is a possibility for $\alpha$, and we know $0 \in W$.
(b) $\beta \neq 0$, then

$$\begin{aligned} \alpha &= \underbrace{a_1\alpha_1 + \cdots + a_v\alpha_v}_{} + \underbrace{a_{v+1}\alpha_{v+1} + \cdots + a_n\alpha_n}_{} \\ &= \qquad \alpha_K \qquad + \qquad \alpha_W \end{aligned}$$

Then

$$\begin{aligned} \sigma(\alpha) &= \sigma(\alpha_K) + \sigma(\alpha_W) \\ &= 0 + \sigma(\alpha_W) \\ &= \sigma(\alpha_W) = \sigma_1(\alpha_W). \end{aligned}$$

So, if instead of $\alpha \in U$ we choose $\alpha_W \in W$ in this way, we find that $\sigma_1(\alpha_W) = \beta$. Thus to every $\beta \in \mathrm{Im}(\sigma)$, there corresponds an element $\alpha_W \in W$. This completes the proof that $\sigma_1$ is an isomorphism of $W$ and $\mathrm{Im}(\sigma)$. Also, since isomorphic vector spaces over the same field have the same dimension we conclude that

$$\dim W = \dim \mathrm{Im}(\sigma)$$

i.e.

$$n - \nu(\sigma) = \rho(\sigma)$$

as required.

Now *READ the first paragraph on page* N32 *beginning* "Theorem 1.6 has an important..." : Then *READ* ***Corollary 1.14*** *on page* N33. You do not need to know a proof of this corollary, but it will be used often when we manipulate matrices. In symbols, it states that if $\tau$ is an isomorphism composable with $\sigma$, then

$$\rho(\tau\sigma) = \rho(\sigma\tau) = \rho(\sigma).$$

**Exercise**

In this exercise we use the linear transformation that we considered in Exercise 4 of page N 35: we repeat its definition here.

Let

$$\sigma : (x_1, x_2, x_3, x_4) \longmapsto (3x_1 - 2x_2 - x_3 - 4x_4, x_1 + x_2 - 2x_3 - 3x_4)$$

(i) Determine $\rho(\sigma)$ and hence find $\nu(\sigma)$.
(ii) Use this information to verify that

$$\{(1, 1, 1, 0), (2, 1, 0, 1)\}$$

is a basis for $K(\sigma)$.

**Solution**

(i) Since $\mathrm{Im}(\sigma) = R^2$, and
$\dim \mathrm{Im}(\sigma) = \rho(\sigma)$, by definition

$$\rho(\sigma) = 2.$$

Hence, by the Dimension Theorem:

$$\dim V = \dim \mathrm{Im}(\sigma) + \dim K(\sigma),$$
$$\dim R^4 = \rho(\sigma) + \nu(\sigma)$$

i.e.

$$4 = 2 + \nu(\sigma)$$

so

$$\nu(\sigma) = 2.$$

(ii) Since $K(\sigma)$ has dimension 2, any two linearly independent vectors in $K(\sigma)$ form a basis for $K(\sigma)$.

The given set

$$\{(1, 1, 1, 0), (2, 1, 0, 1)\}$$

has two members (both in $K(\sigma)$), and is linearly independent, since

$$a(1, 1, 1, 0) + b(2, 1, 0, 1) = (0, 0, 0, 0)$$

implies $a = b = 0$.

So the given set is a basis for $K(\sigma)$.

## 2.1.8 Specifying Linear Transformations

So far we have specified our linear transformations by specifying their effect on every vector in their domains. It is sufficient, however, to specify their effect only on a basis in the domain. The next theorem explains why.

*READ **Theorem 1.17** on page* N34 *and its proof.*

This theorem will be used in *Unit 3, Hermite Normal Form*, where we shall find it useful, given an arbitrary set of $n$ vectors in a space $V$, to think of them as the images of the basis vectors in the domain $U$ of a linear transformation $\sigma$. The linear transformation obtained depends on what domain $U$ is chosen (any $n$-dimensional vector space will do) and what basis is chosen in $U$. We include the proof as an exercise.

**Exercise**

Fill in the missing words or symbols in the following proof of ***Theorem 1.17***.
Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis for $U$. Then any vector $\alpha \in U$ can be expressed uniquely in the form

$$\alpha = \boxed{\phantom{xxxxxxxxxxxxxxxx}} \qquad \text{(i)}$$

where $a_1, \ldots, a_n$ are scalars, called the

$$\boxed{\phantom{xxxxxxxxxxxxxxxx}} \qquad \text{(ii)}$$

of $\alpha$ with respect to the basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

We define the mapping

$$\sigma : \alpha \longmapsto a_1\beta_1 + \cdots + a_n\beta_n \qquad (\alpha \in U)$$

where $\beta_i \in V$, $i = 1, 2, \ldots, n$.

Since the representation (i) is unique, $\sigma$ is a

$\boxed{\qquad\qquad\qquad}$ (iii)

and it has the required property $\sigma(\alpha_i) = \beta_i$ $(i = 1, 2, \ldots, n)$, but we do not yet know whether it is a linear transformation, or whether this linear transformation is unique. To prove the linearity, let $c, c' \in F$ and $\alpha, \alpha' \in U$.

Then, if

$$\alpha = a_1\alpha_1 + \cdots + a_n\alpha_n$$

and

$$\alpha' = a'\alpha_1 + \cdots + a'_n\alpha_n,$$

we have the unique representation (as in (i))

$$c\alpha + c'\alpha' = (ca_1 + c'a'_1)\alpha_1 + \cdots + (ca_n + c'a'_n)\alpha_n$$

so that

$$\sigma(c\alpha + c'\alpha') = \boxed{\qquad\qquad\qquad\qquad\qquad} \quad \text{(iv)}$$

$$= \boxed{\qquad\qquad\qquad\qquad\qquad} \quad \text{(v)}$$

which proves that $\sigma$ is a linear transformation. To prove that $\sigma$ is the only linear transformation with the required property, we show that any linear transformation $\tau$ with this property

$$\tau(\alpha_i) = \beta_i \qquad (i = 1, \ldots, n)$$

is in fact $\sigma$. We know already that

$$\sigma(\alpha_i) = \boxed{\qquad\qquad} \qquad (i = 1, \ldots, n) \qquad \text{(vi)}$$

It follows that

$$(\sigma - \tau)(\alpha_i) = \boxed{\qquad\qquad} \qquad (i = 1, \ldots, n) \qquad \text{(vii)}$$

and, since every element of $U$ is a linear combination of $\{\alpha_1, \ldots, \alpha_n\}$, that $\sigma - \tau$ is the

$\boxed{\qquad\qquad}$ (viii)

transformation. Thus we have shown that $\tau = \sigma$; i.e. $\sigma$ is the only linear transformation with the required property.

**Solution**

(i) $a_1\alpha_1 + \cdots + a_n\alpha_n$
(ii) coordinates
(iii) function
(iv) $(ca_1 + c'a'_1)\beta_1 + \cdots + (ca_n + c'a'_n)\beta_n$
(v) $c\sigma(\alpha) + c'\sigma(\alpha')$.
(vi) $\beta_i$
(vii) 0
(viii) zero

Finally *READ from the end of* ***Theorem 1.17*** *on page* N34 *to the end of the first paragraph at the top of page* N35.

We use the corollary in place of ***Theorem 1.17*** when faced with less than $n$ vectors in $V$.

### 2.1.9 Summary of Section 2.1

In this section we defined the terms

| | | |
|---|---|---|
| linear transformation | (page N27) | * * * |
| image of a linear transformation | (page N27) | * * * |
| inverse image | (page N27) | * * |
| complete inverse image | (page N27) | * * |
| homomorphism | (page N27) | * * |
| one–one | (page N27) | * * * |
| onto | (page N28) | * * * |
| isomorphism | (page N28) | * * * |
| rank of a linear transformation | (page N31) | * * * |
| kernel of a linear transformation | (page N31) | * * * |
| nullity of a linear transformation | (page N31) | * * |

We introduced the notation

| | | |
|---|---|---|
| $\mathrm{Im}(\sigma)$ : | the image of a linear transformation, $\sigma$ | (page N28) |
| $\rho(\sigma)$ : | the rank of a linear transformation, $\sigma$ | (page N31) |
| $K(\sigma)$ : | the kernel of a linear transformation, $\sigma$ | (page N31) |
| $\nu(\sigma)$ : | the nullity of a linear transformation, $\sigma$ | (page N31) |

**Theorems**

1. (*1.1*, page N28)
   The inverse $\sigma^{-1}$ of an isomorphism is also an isomorphism. * * *
2. (page N29)
   If $V$ is a vector space over $F$ and dim $V = n$, then $V$ is isomorphic to $F^n$. * * *
3. (*1.2*, page N31)
   $\mathrm{Im}(\sigma)$ is a subspace of $V$ (the codomain of $\sigma$). * * *
4. (page N31)
   $K(\sigma)$ is a subspace of $U$ (the domain of $\sigma$). * * *
5. (*1.6*, page N31)
   $\rho(\sigma) + \nu(\sigma) = n$ (the dimension of the domain of $\sigma$). * * *
6. (*1.14*, page N33)
   $\rho(\sigma) = \rho(\sigma\tau_1) = \rho(\tau_2\sigma)$ if $\tau_1$, $\tau_2$ are isomorphisms. * *
7. (*1.17*, page N34)
   Let $A = \{\alpha_1, \ldots, \alpha_n\}$ be any basis of $U$. * * *
   Let $B = \{\beta_1, \ldots, \beta_n\}$ be any $n$ vectors in $V$ (not necessarily linearly independent). There exists a uniquely determined transformation $\sigma$ of $U$ into $V$ such that $\sigma(\alpha_i) = \beta_i$ for $i = 1, 2, \ldots, n$.

**Techniques**

1. Decide whether a given mapping is a linear transformation. * * *
2. Decide whether a given linear transformation $\sigma$ is * * *
   (a) one–one
   (b) onto
   (c) an isomorphism.
3. Given an isomorphism $\sigma$, determine $\sigma^{-1}$. * *
4. Find $\sigma_1 + \sigma_2$, $\sigma_1\sigma_2$, $\sigma_2\sigma_1$ (if they exist), given $\sigma_1$ and $\sigma_2$. *
5. Given $\sigma$ and a basis for its domain, determine $K(\sigma)$. * * *
6. Find $\rho(\sigma)$, $\nu(\sigma)$, $\mathrm{Im}(\sigma)$ for a given $\sigma$. * *

## 2.2 MATRICES

### 2.2.1 Isomorphisms Between Linear Transformations and Matrices

Section 2 should be rather easy going for you, as there is only one new idea. This is the fact that there is an isomorphism between linear transformations and matrices. Putting the horse before the cart, so to speak, we shall say that the matrix *faithfully represents* the linear transformation. Every theorem that is true for linear transformations is true for the representing matrices, and vice versa! Just as with vectors this numerical representation is essential for most computations involving linear transformations, and, just as with vectors, the actual numbers appearing in the representation depend on the choice of basis.

*READ from the beginning of Section 2 on page* N37 *to Equation (2.4) on page* N39.

#### Notes

(i) *Equation (2.2), page* N38 This equation is one of the most important equations that we shall meet in this course. Let us examine it. Suppose $[a_{ij}] = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$; then Equation 2.2 tells us that

$$\sigma(\alpha_1) = a_{11}\beta_1 + a_{21}\beta_2 = \beta_1 + 3\beta_2$$
$$\sigma(\alpha_2) = a_{12}\beta_1 + a_{22}\beta_2 = 2\beta_1 + 4\beta_2 .$$

In words : *each column of the matrix* $[a_{ij}]$ *gives the coordinates of the image of one of the domain basis vectors.* Here the first column of $[a_{ij}]$ is $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\sigma(\alpha_1) = 1\beta_1 + 3\beta_2$ ; the second column of $[a_{ij}]$ is $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ and $\sigma(\alpha_2) = 2\beta_1 + 4\beta_2$.

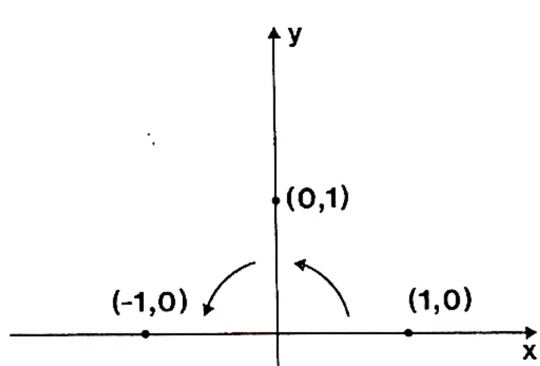
We often use this to advantage. For if we are explicitly given the images of a basis of $U$ as linear combinations of basis vectors of $V$, we just arrange the coefficients of each equation in columns to construct the representing matrix.

(ii) *Line 2 after Equation (2.2) on page* N38 Be sure not to confuse the type-faces. The sans-serif $\mathsf{A}$, representing a basis in $U$, has nothing to do with the italic $A$, representing the matrix.

(iii) *line 3, page* N39 The rotations mentioned are illustrated below. The fact that rotations about the origin are linear transformations greatly simplifies the calculation of the formulas for rotations: we have only to compute the effect of the rotation on a basis and the whole transformation is determined (because of *Theorem 1.17* on page N34).

*Example*

90° rotation anticlockwise about the origin

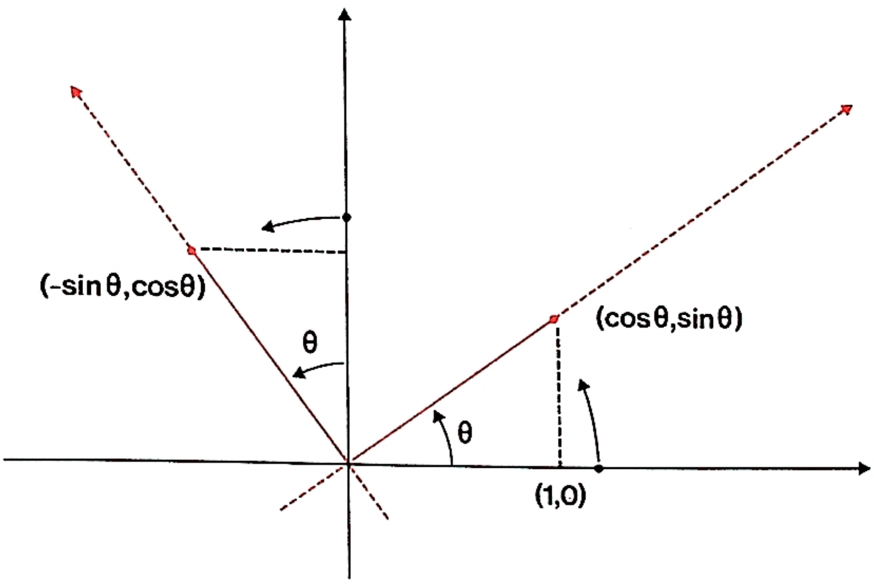



$$\left.\begin{array}{l} (1, 0) \longmapsto (0, 1) \\ (0, 1) \longmapsto (-1, 0) \end{array}\right\} \text{ coordinates of images of basis}$$

Putting the two coordinates (0, 1) and $(-1, 0)$ into the first and second columns, we get the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ given in N.

*Example*

General rotation of $\theta$ about the origin




$$(1, 0) \longmapsto (\cos\theta, \sin\theta)$$
$$(0, 1) \longmapsto (-\sin\theta, \cos\theta)$$

The first column of the matrix is therefore

$$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \quad \text{the second is} \quad \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix},$$

and the complete matrix is given by (2.4) of N. If you like geometric illustrations of this type, you may like to try Exercises 7 and 8 of page N43.

**Exercises**

1. Exercise 4, page N42.
2. Exercise 5, page N43.
3. Exercise 6, page N43.

**Solutions**

1. Let $\alpha_1 = (1, 0)$ and $\alpha_2 = (0, 1)$ denote the given basis.

   Then we have $\sigma(\alpha_1) = (3, -1) = 3\alpha_1 - \alpha_2$

   $$\sigma(\alpha_2) = (-1, 2) = -\alpha_1 + 2\alpha_2$$

   Filling up the columns of the matrix as explained in note (i) to the previous reading passage, we obtain the matrix

   $$\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}.$$

2. With $\alpha_1 = (1, 0)$ and $\alpha_2 = (0, 1)$ we have, using N's hint,

   $$\begin{aligned} \sigma((1, 0)) &= \sigma(\tfrac{1}{2}(1, 1) + \tfrac{1}{2}(1, -1)) \\ &= \tfrac{1}{2}\sigma((1, 1)) + \tfrac{1}{2}\sigma((1, -1)) \\ &= \tfrac{1}{2}(2, -3) + \tfrac{1}{2}(4, -7) \\ &= (3, -5) = 3\alpha_1 - 5\alpha_2. \end{aligned}$$

   Similarly

   $$\begin{aligned} \sigma((0, 1)) &= \sigma(\tfrac{1}{2}(1, 1) - \tfrac{1}{2}(1, -1)) \\ &= \tfrac{1}{2}(2, -3) - \tfrac{1}{2}(4, -7) \\ &= (-1, 2) = -\alpha_1 + 2\alpha_2. \end{aligned}$$

Filling in the columns as before, we get the matrix

$$\begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}.$$

3. We use $(1, 0) = a(3, -1) + b(-1, 2)$ to find $a$, $b$. This gives

$$(1, 0) = \tfrac{2}{5}(3, -1) + \tfrac{1}{5}(-1, 2).$$

Similarly, we find

$$(0, 1) = \tfrac{1}{5}(3, -1) + \tfrac{3}{5}(-1, 2).$$

Then $\sigma^{-1}((1, 0)) = \frac{2}{5}\sigma^{-1}((3, -1)) + \frac{1}{5}\sigma^{-1}((-1, 2))$ and this gives the image of a basis element in terms of basis elements, so the first column of the matrix representing $\sigma^{-1}$ is

$$\begin{pmatrix} \frac{2}{5} \\ \frac{1}{5} \end{pmatrix}$$

Also

$$\sigma^{-1}((0, 1)) = \tfrac{1}{5}\sigma^{-1}((3, -1)) + \tfrac{3}{5}\sigma^{-1}((-1, 2)),$$

which gives the second column of the matrix; and so the complete matrix is

$$\begin{bmatrix} \frac{2}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{3}{5} \end{bmatrix}.$$

### 2.2.2 Matrix Algebra

Since matrices represent linear transformations, the operations of scalar multiplication, addition, and multiplication which we can perform with linear transformations have their counterparts for matrices. These are described in the next reading passage, which is essentially a generalization of material already treated in *Unit M100 23.*

*READ from the line following Equation (2.4) on page* N39 to *page* N41, *line 1.*

**Exercises**

1. Exercise 1, page N42.
2. Exercise 2, page N42.
3. Exercise 3, page N42.

**Solutions**

The answers are given on pages N326–327. Note in Exercise 2 that the result could have been obtained by using the products in Exercise 1. Write products of 1(*a*), (*b*) and (*c*) as

$$\begin{aligned} AB &= C \qquad (a) \\ BD &= E \qquad (b) \\ AE &= F \qquad (c) \end{aligned}$$

From (*b*)

$$ABD = AE$$

But from (*a*) $AB = C$

and from (*c*) $AE = F$

hence $\quad CD = F.$

This is the result you should have calculated. Note the use of property 5 on page N40, in the form

$$A(BD) = (AB)D.$$

Note in Exercise 3 that $AB \neq BA$: in general, matrix multiplication is not commutative. This corresponds to the fact that, in general, the product of linear transformations is not commutative (see page N30). Also ignore N's comment about the rank of $AB$ and $BA$ (*we* haven't defined the rank of a matrix yet).

## 2.2.3 Rank and Nullity: Simultaneous Equations

The last reading passage in this section deals with two topics; first, how we define rank for matrices, and second, how we use matrices to write systems of simultaneous linear equations in a concise form.

*READ from line 2 of page* N41 *to the end of the section, line 7, page* N42.

**Notes**

(i) ***Theorem 2.1*** The first statement in this theorem is the statement of the dimension theorem in terms of matrices. The second statement is the statement of ***Theorem 1.12*** on page N33 expressed in matrix form. We omitted ***Theorem 1.12***, but you should note the result here in terms of matrices.
(ii) *line 11, page* N41 A maximal linearly independent subset of $\{\sigma(\alpha_1), \ldots, \sigma(\alpha_n)\}$ is a linearly independent subset which is not contained in a larger linearly independent subset. Thus in the subset of $R^2$

$$\{(1, 0), (0, 1), (1, -1), (\tfrac{1}{2}, 0)\},$$

$\{(1, 0), (0, 1)\}$, $\{(1, 0), (1, -1)\}$, etc, are maximal linearly independent subsets. (This term is defined on page N14, Exercise 3.)
(iii) *Equation (2.8), page* N41 Notice that the sum is now over the second subscript, which labels the columns of the matrix, not the first, which labels the rows, as in Equation (2.2). Whenever we multiply two matrices together, the subscripts to be summed over are contiguous and equal (as for the subscript "$i$" in the law of matrix multiplication,

$$c_{kj} = \textstyle\sum_i b_{ki} a_{ij}).$$

In Equation (2.2) on page N38, on the other hand, we were not multiplying two matrices together: the $\beta$s are a set of vectors, not elements of a matrix.
(iv) *line −4, page* N41 "Matric" is an adjectival form of the word "matrix".
(v) *line 2, page* N42 We now have four different ways of representing a vector $\xi$ belonging to an $n$-dimensional space $U$:

(a) the $n$-tuple $(x_1, \ldots, x_n) \in F^n$

(b) the one-column matrix $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

(c) the one-row matrix $[x_1\ x_2\ \cdots\ x_n]$

(d) the symbol $(x_1, \ldots, x_n)$ used for typographical convenience as abbreviation for the column matrix shown in (b).

Use whichever of the four representations you like—they are all connected by isomorphisms—but do not confuse (a) and (d).

### 2.2.4 Summary of Section 2.2

In this section we defined the terms

| | | |
|---|---|---|
| matrix | (page N37) | * * * |
| order of a matrix | (page N37) | * * * |
| row, column of a matrix | (page N37) | * * * |
| matrix representing a linear transformation | (page N38) | * * * |
| rank and nullity of a matrix | (page N41) | * * * |

**Theorem**

(*2.1*, page N41)
For an $m \times n$ matrix $A$, the rank of $A$ plus the nullity of $A$ is equal to $n$. The rank of a product $BA$ is less than or equal to the rank of either factor. * * *

**Techniques**

1. Determine the matrix representing a linear transformation with respect to given bases of the domain and of the codomain (Equation (2.2), page N38). * * *
2. Add matrices (Equation (2.5), page N39) * * *
   Multiply a matrix by a scalar (Equation (2.6), page N39) * * *
   Multiply matrices (Equation (2.7), page N40) * * *
3. Represent vectors and $n$-tuples (vectors in $F^n$) by both row and column matrices (pages N41 and 42). * *

## 2.3 NON-SINGULAR MATRICES

### 2.3.1 Non-singular Matrices

This short section introduces a classification among square matrices representing linear transformations $U \longrightarrow U$. The classification is based on this question: given a square matrix $A$, does there exist a matrix $B$ such that $AB = I$, where $I$ is the square matrix

$$\begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix}$$

with the same number of rows and columns as $A$? The matrix $I$ represents the identity transformation $U \longrightarrow U$. If $B$ does exist, $A$ is said to be *non-singular* or *invertible* and we write $B = A^{-1}$. It also follows that $A^{-1}A = I$.

Otherwise, $A$ is said to be *singular*.

For example, the $2 \times 2$ matrices $R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, defined in the previous section to represent rotations about the origin, are all non-singular, because

$$R(\theta)R(-\theta) = R(\theta - \theta) = R(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

On the other hand, the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is singular because the matrix equation

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

has no solution (work out the product on the left if you can't see why).

*READ the whole of Section 3, pages* N45–N48. (*You have read the last paragraph already.*)

### Notes

(i) *Line 7, page* N46 The symbol $\delta_{ij}$ is the so-called Kronecker Delta, which we first met on page N15.
(ii) ***Theorem 3.1*** This is a statement of ***Theorem 1.1*** (page N28) for automorphisms.
(iii) ***Theorem 3.2*** Ignore the words "*that is, if and only if it is an epimorphism*".
(iv) ***Theorem 3.3*** Ignore the words "*that is, if and only if it is a monomorphism*".

These two theorems are a restatement of ***Theorem 1.8.*** (page N32), which we omitted previously. Both theorems can be derived directly from the dimension theorem: we illustrate this by proving ***Theorem 3.2***. (The proofs are not important, but you should read the one given: this sort of proof occurs frequently in linear algebra.) We have to prove two things (notice the "*if and only if*" !)

(a) if $\tau$ is an automorphism, then it is of rank $n$;
(b) if $\tau$ is of rank $n$, then it is an automorphism.

*Proof*

(a) $\tau$ is an automorphism, so it is one-to-one and onto. Since it is onto, $\mathrm{Im}(\tau) = V$, so that $\rho(\tau) = n$.
(b) $\tau$ is of rank $n$, i.e. $\rho(\tau) = n$. This means that $\mathrm{Im}(\tau)$ is of dimension $n$ and $\tau$ is therefore "onto". To show that it is also one-to-one we use the dimension theorem. This tells us that the kernel of $\tau$ has dimension 0, since $\rho(\tau) = \dim \mathrm{Im}(\tau) = n$, and hence that

$$\tau(x) = \tau(y)$$

implies

$$x = y,$$

since

$\tau(x-y)=0$ implies $x-y=0$.

(iv) *line −10 to −9, page* N46 The statement "automorphisms are the only linear transformations which have inverses" is not correct: *every* isomorphism has an inverse. What N means is "automorphisms are the only endomorphisms (linear transformations of a set to itself) which have inverses".
(v) *line 2 et seq., page* N47 The deduction in line 2 is an application of the second part of ***Theorem 2.1***. You should understand the proofs given. These theorems are often used in manipulating matrices.

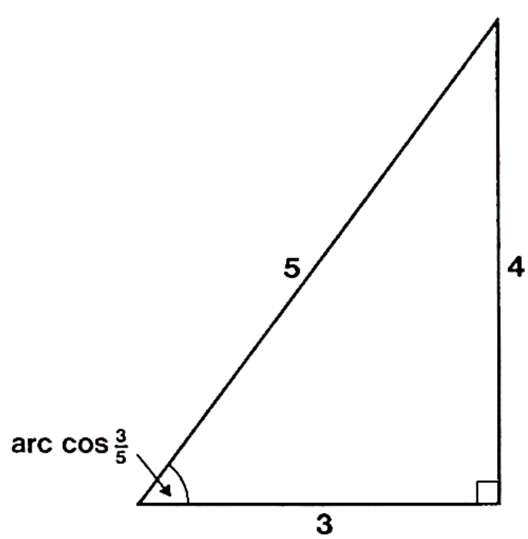
Notice particularly that in taking the inverse of a matrix product, the order of the factors must be reversed:

$$(AB)^{-1}=B^{-1}A^{-1}$$

(vi) ***Theorem 3.6*** This theorem is a little easier to understand if put this way: Let $A$ be non-singular. Then $AY=B$ has solution $Y=A^{-1}B$ and $XA=B$ has solution $X=BA^{-1}$. These solutions are unique. It is not generally true that $A^{-1}B=BA^{-1}$.
(vii) ***Theorem 3.7*** This theorem is a direct consequence of ***Corollary 1.14*** (*page* N33), and you should know the result and be able to recognize its implicit use later on in the course. No proof is necessary.

**Exercises**

1. Exercise 1, page N48.
2. Exercise 2, page N48. ("Find the square" means multiply the matrix by itself.)
3. Exercise 3, page N48.
4. Exercise 5, page N49. (In this question, arc cos $\frac{3}{5}$ means "the angle in the first quadrant whose cosine is $\frac{3}{5}$. The sine of this angle is $\frac{4}{5}$.)




5. Exercise 10, page N49.

**Solutions**

1. Show that $A^{-1}A=I$ or that $AA^{-1}=I$.

2. $$A^2=\frac{1}{9}\begin{bmatrix}9&0&0\\0&9&0\\0&0&9\end{bmatrix}=I.\ \text{Thus } A^{-1}=A.$$

Note that $A^2=I$ does not imply $A=I$.

3. $$\begin{bmatrix}1&2&3\\2&3&4\\0&1&2\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix}=\begin{bmatrix}1-4+3\\2-6+4\\-2+2\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

If there were an $A^{-1}$, we would have

$$\begin{aligned}A^{-1}AX&=A^{-1}0\\&=0.\end{aligned}$$

But

$$A^{-1}AX = X$$

so

$$X = 0$$

But

$$X \neq 0$$

so $A^{-1}$ does not exist.

Alternatively, since $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \longmapsto \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, the kernel of the transformation has positive dimension, so that the transformation is not one-to-one and therefore not invertible.

4. The given matrix is of the form $R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. The inverse rotation is obtained by rotating back through $\theta$, i.e. through $-\theta$, so that the inverse matrix is

$$\begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos\theta & +\sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$
$$= \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix}.$$

The product $\begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

5. A sufficient clue to all but the last part of this exercise is given in the answer on page N328. For the last part, suppose that $n > m$, then the rank of $BA$ cannot exceed $m$ (*Theorem 2.1*), and so is less than $n$. So $BA$ cannot be the $n \times n$ identity matrix which is of rank $n$.

## 2.3.2 Summary of Section 2.3

In this section we defined the terms

| | | |
|---|---|---|
| endomorphism | (page N45) | * * * |
| non-singular or invertible matrix | (page N46) | * * * |
| singular matrix | (page N46) | * * * |
| automorphism | (page N46) | * |

**Theorems**

1. (*3.6*, page N47)
If $A$ is non-singular, we can solve uniquely the equations $XA = B$ and $AY = B$ for any matrix $B$ of the proper size, but the two solutions need not be equal. * * *
2. (*3.7*, page N48)
The rank of a (not necessarily square) matrix is not changed by multiplication by a non-singular matrix. * * *

**Techniques**

Use the following results when manipulating matrices:

(i) $A^{-1}A = AA^{-1} = I$ * * *
(ii) $(aA)^{-1} = a^{-1}A^{-1}$, $a \in F$ * * *
(iii) $(AB)^{-1} = B^{-1}A^{-1}$ * * *

## 2.4 SUMMARY OF THE UNIT

### Definitions

The terms defined in this unit and page references to their definitions are given below.

| | | |
|---|---|---|
| linear transformation | (page N27) | * * * |
| image of a linear transformation | (page N27) | * * * |
| inverse image | (page N27) | * * |
| complete inverse image | (page N27) | * * |
| homomorphism | (page N27) | * * |
| one–one | (page N27) | * * * |
| onto | (page N28) | * * * |
| isomorphism | (page N28) | * * * |
| rank of a linear transformation | (page N31) | * * * |
| kernel of a linear transformation | (page N31) | * * * |
| nullity of a linear transformation | (page N31) | * * |
| matrix | (page N37) | * * * |
| order of a matrix | (page N37) | * * * |
| row, column of a matrix | (page N37) | * * * |
| matrix representing a linear transformation | (page N38) | * * * |
| rank of a matrix | (page N41) | * * * |
| nullity of a matrix | (page N41) | * * * |
| endomorphism | (page N45) | * * * |
| non-singular or invertible matrix | (page N46) | * * * |
| singular matrix | (page N46) | * * * |
| automorphism | (page N46) | * |

### Theorems

We list the important theorems discussed in this unit. Only 2- and 3-star theorems which are essential to this and later units have been included in this list. References to the statement of the theorems in N are also given.

1. (*1.1*, page N28)
   The inverse $\sigma^{-1}$ of an isomorphism is also an isomorphism. * * *
2. (page N29)
   If $V$ is a vector space over $F$ and dim $V = n$, then $V$ is isomorphic to $F^n$. * * *
3. (*1.2*, page N31)
   $\mathrm{Im}(\sigma)$ is a subspace of $V$ (the codomain of $\sigma$). * * *
4. (page N31)
   $K(\sigma)$ is a subspace of $U$ (the domain of $\sigma$). * * *
5. (*1.6*, page N31)
   $\rho(\sigma) + \nu(\sigma) = n$ (the dimension of the domain of $\sigma$). * * *
6. (*1.14*, page N33)
   $\rho(\sigma) = \rho(\sigma\tau_1) = \rho(\tau_2\sigma)$ if $\tau_1$, $\tau_2$ are isomorphisms. * *
7. (*1.17*, page N34)
   Let $A = \{\alpha_1, \ldots, \alpha_n\}$ be any basis of $U$. * * *
   Let $B = \{\beta_1, \ldots, \beta_n\}$ be any $n$ vectors in $V$ (not necessarily linearly independent). There exists a uniquely determined transformation $\sigma$ of $U$ into $V$ such that $\sigma(\alpha_i) = \beta_i$ for $i = 1, 2, \ldots, n$.
8. (*2.1*, page N41)
   For an $m \times n$ matrix $A$, the rank of $A$ plus the nullity of $A$ is equal to $n$. * * *
   The rank of a product $BA$ is less than or equal to the rank of either factor.
9. (*3.6*, page N47)
   If $A$ is non-singular, we can solve uniquely the equations $XA = B$ and $AY = B$ for any matrix $B$ of the proper size, but the two solutions need not be equal. * * *
10. (*3.7*, page N48)
    The rank of a (not necessarily square) matrix is not changed by multiplication by a non-singular matrix. * * *

**Techniques**

1. Decide whether a given $\sigma$ is a linear transformation. * * *
2. Decide whether a given linear transformation is * * *

   (a) one–one
   (b) onto
   (c) an isomorphism.

3. Given an isomorphism $\sigma$, determine $\sigma^{-1}$. * *
4. Find $\sigma_1 + \sigma_2$, $\sigma_1\sigma_2$, $\sigma_2\sigma_1$ (if they exist), given $\sigma_1$ and $\sigma_2$. *
5. Given $\sigma$ and a basis for its domain, determine $K(\sigma)$. * * *
6. Find $\rho(\sigma)$, $\nu(\sigma)$, $\mathrm{Im}(\sigma)$ for a given $\sigma$. * *
7. Determine the matrix representing a linear transformation with respect to given bases of the domain and of the codomain (Equation (2.2), page N38). * * *
8. Add matrices (Equation (2.5), page N39) * * *
   Multiply a matrix by a scalar (Equation (2.6), page N39) * * *
   Multiply matrices (Equation (2.7), page N40) * * *
9. Represent vectors and $n$-tuples (vectors in $F^n$) by both row and column matrices (pages N41 and 42). * *
10. Use the following results when manipulating matrices: * * *

    (i) $A^{-1}A = AA^{-1} = I$
    (ii) $(aA)^{-1} = a^{-1}A^{-1}$, $a \in F$
    (iii) $(AB)^{-1} = B^{-1}A^{-1}$.

**Notation**

$\mathrm{Im}(\sigma)$ : the image of a linear transformation, $\sigma$ (page N28)
$\rho(\sigma)$ : the rank of a linear transformation, $\sigma$ (page N31)
$K(\sigma)$ : the kernel of a linear transformation, $\sigma$ (page N31)
$\nu(\sigma)$ : the nullity of a linear transformation, $\sigma$ (page N31)

## 2.5 SELF-ASSESSMENT

### Self-assessment Test

This Self-assessment Test is designed to help you test quickly your understanding of the unit. It can also be used, together with the summary of the unit, for revision. The answers to these questions will be found on the next non-facing page. We suggest you complete the whole test before looking at the answers.

1. Let $\sigma$ be any linear transformation, $\sigma : U \longrightarrow V$. Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis of $U$, and let $\alpha = 2\alpha_1 - \alpha_2 + \alpha_n$. Express $\sigma(\alpha)$ as a linear combination of $\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_n)$.
2. Complete the following statement. Let $\sigma$ be a transformation of $U$ into $V$. In order to prove that $\alpha$ is a *linear transformation*, we take arbitrary $\alpha, \beta \in U$, arbitrary $a, b \in F$ and show that ____________________.
3. Let $\sigma : R^2 \longrightarrow R^2$ be $\sigma : (x_1, x_2) \longmapsto (x_1 + x_2, x_2)$. Is $\sigma$ a linear transformation?
4. Which of the following mappings are isomorphisms?

   (a) $U = R^3,\ V = R^3;\ \sigma : (x_1, x_2, x_3) \longmapsto (x_3, x_2, x_1)$
   (b) $U = R^3,\ V = R^3;\ \sigma : (x_1, x_2, x_3) \longmapsto (x_1, x_2, x_2)$
   (c) $U = R^3,\ V = P_3;\ \sigma : (x_1, x_2, x_3) \longmapsto x_1 + x_2 t + x_3 t^2$
   (d) $U = R^2,\ V = R^3;\ \sigma : (x_1, x_2) \longmapsto (0, x_1, x_2)$.

5. Let $\sigma : R^3 \longmapsto R^3$ be

   $$\sigma : (x_1, x_2, x_3) \longmapsto (x_1 - x_2, x_2 - x_3, x_3 - x_1).$$

   (i) What is the kernel of $\sigma$?
   (ii) What is the rank of $\sigma$?

6. Derive the matrix representing the automorphism of $R^3$ defined by:

   $$(x_1, x_2, x_3) \longmapsto (x_3, -x_1, x_2)$$

   with respect to the basis $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$, for both domain and codomain.

7. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$

   Determine $AD - DA$.
8. If $A$, $B$ are non-singular $n \times n$ matrices, which of the following is equal to $(B^{-1}A)^{-1}$?

   (a) $BA$ (b) $A^{-1}B$ (c) $B^{-1}A$
   (d) $A^{-1}B^{-1}$ (e) $BA^{-1}$ (f) $AB^{-1}$.

## Solutions to Self-assessment Test

**1.** $$\sigma(\alpha) = \sigma(2\alpha_1 - \alpha_2 + \alpha_n)$$

Since $\sigma$ is a linear transformation,

$$\begin{aligned}\sigma(\alpha) &= \sigma(2\alpha_1) + \sigma(-\alpha_2) + \sigma(\alpha_n)\\ &= 2\sigma(\alpha_1) - \sigma(\alpha_2) + \sigma(\alpha_n).\end{aligned}$$

**2.** $$\cdots \sigma(a\alpha + b\beta) = a\sigma(\alpha) + b\sigma(\beta)$$

(An equivalent method is to show that

$$\sigma(a\alpha) = a\sigma(\alpha)$$

*and*

$$\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta).)$$

**3.** Yes.

$$\begin{aligned}&\sigma(a(x_1, x_2) + b(y_1, y_2))\\ &\quad = \sigma((ax_1 + by_1, ax_2 + by_2))\\ &\quad = (ax_1 + by_1 + ax_2 + by_2, ax_2 + by_2)\\ &\quad = (a(x_1 + x_2) + b(y_1 + y_2), ax_2 + by_2)\\ &\quad = a(x_1 + x_2, x_2) + b(y_1 + y_2, y_2)\\ &\quad = a\sigma(x_1, x_2) + b\sigma(y_1, y_2).\end{aligned}$$

**4.** (a) and (c) are isomorphisms.
(b) and (d) are not isomorphisms. For (b), we see that $\sigma$ is not onto (for example, $(1, 2, 3) \notin \mathrm{Im}(\sigma)$) nor is it one–one (for example, $\sigma((1, 2, 3)) = \sigma((1, 2, 4)) = (1, 2, 2)$).
For (d), the mapping is not onto (for example, $(1, 2, 3) \notin \mathrm{Im}(\sigma)$). $\sigma$ is one–one, however.

**5.** (i) If $(x_1, x_2, x_3) \in K(\sigma)$

$$\sigma((x_1, x_2, x_3)) = (0, 0, 0)$$

i.e.

$$(x_1 - x_2, x_2 - x_3, x_3 - x_1) = (0, 0, 0).$$

So

$$\begin{aligned}x_1 &= x_2\\ x_2 &= x_3\\ x_3 &= x_1\end{aligned}$$

i.e.

$$x_1 = x_2 = x_3$$

and hence $K(\sigma) = \{(x_1, x_2, x_3) : x_1 = x_2 = x_3 = a, a \in R\}$.

(ii) $$\rho(\sigma) = \dim \mathrm{Im}(\sigma)$$

If $(x_1, x_2, x_3) \in \mathrm{Im}(\sigma)$,

$$x_1 + x_2 + x_3 = 0$$

thus $\mathrm{Im}(\sigma) = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 0\}$
$\qquad = \{(x_1, x_2, -(x_1 + x_2))\}$

i.e.

$$\dim \mathrm{Im}(\sigma) = \rho(\sigma) = 2.$$

*Alternatively*, by the dimension theorem

$$n = \nu + \rho$$

i.e.

$$\begin{aligned}\rho &= n - \nu \\ &= 3 - 1 \\ &= 2\end{aligned}$$

since

$$\nu = \dim K(\sigma) = 1.$$

6. $\sigma : (x_1, x_2, x_3) \longmapsto (x_3, -x_1, x_2)$

So

$$\begin{aligned}\sigma((1, 0, 0)) &= (0, -1, 0)\\ \sigma((0, 1, 0)) &= (0, 0, 1)\\ \sigma((0, 0, 1)) &= (1, 0, 0).\end{aligned}$$

Using Equation (2.2) on page N38, viz

$$\sigma(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_i,$$

$$\begin{aligned}\sigma((1, 0, 0)) &= -(0, 1, 0)\\ &= 0(1, 0, 0) - 1(0, 1, 0) + 0(0, 0, 1)\\ \sigma((0, 1, 0)) &= (0, 0, 1)\\ &= 0(1, 0, 0) + 0(0, 1, 0) + 1(0, 0, 1)\\ \sigma((0, 0, 1)) &= (1, 0, 0)\\ &= 1(1, 0, 0) + 0(0, 1, 0) + 0(0, 0, 1).\end{aligned}$$

Hence the matrix representing the transformation is

$$\begin{bmatrix} 0 & 0 & 1\\ -1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}.$$

7. $AD = \begin{bmatrix} a & b\\ c & d\end{bmatrix}\begin{bmatrix} 2 & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix} a \times 2 + 0 & 0 + b \times 1\\ c \times 2 + 0 & 0 + d \times 1\end{bmatrix}$

$$= \begin{bmatrix} 2a & b\\ 2c & d\end{bmatrix}.$$

$DA = \begin{bmatrix} 2 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix} a & b\\ c & d\end{bmatrix} = \begin{bmatrix} 2 \times a + 0 & 2 \times b + 0\\ 0 + 1 \times c & 0 + 1 \times d\end{bmatrix}$

$$= \begin{bmatrix} 2a & 2b\\ c & d\end{bmatrix}.$$

Hence $AD - DA = \begin{bmatrix} 0 & -b\\ c & 0\end{bmatrix}.$

8. Since $(AB)^{-1} = B^{-1}A^{-1}$

$$\begin{aligned}(B^{-1}A)^{-1} &= A^{-1}(B^{-1})^{-1}\\ &= A^{-1}B.\end{aligned}$$

Hence (b) is the correct answer.

335 01090 3
335 01091 1